

## ПРИМЕНЕНИЕ ПЛИС ДЛЯ СОЗДАНИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ИХ КОМПОНЕНТОВ

\*Институт кибернетики им. В.М. Глушкова НАН Украины, Киев, Украина

**Анотація.** Розглянуто особливості організації архітектури супер-ЕОМ, виконаних на ПЛИС, запропоновано основні принципи їх побудови і приведені особливості архітектурно-структурної організації реконфігурованих обчислювальних систем (РОС), зокрема: на основі базових модулів, виконаних на ПЛИС; РОС з відкритою масштабованою архітектурою (РОС ОМА); з безпосередніми зв'язками між обчислювальними модулями і порозрядною обробкою операндів. Приведено приклад застосування ПЛИС для побудови гетерогенної супер-ЕОМ «СКІФ» ряду 4.

**Ключові слова:** застосування ПЛИС, реконфігурована обчислювальна система, архітектурно-структурна організація супер-ЕОМ, базовий модуль на ПЛИС, відкрита масштабована архітектура.

**Аннотация.** Рассмотрены особенности организации архитектуры супер-ЭВМ, выполненных на ПЛИС, предложены основные принципы их построения и приведены особенности архитектурно-структурной организации реконфигурируемых вычислительных систем (РВС), в том числе: на основе базовых модулей, выполненных на ПЛИС; РВС с открытой масштабируемой архитектурой (РВС ОМА); с непосредственными связями между вычислительными модулями и поразрядной обработкой операндов. Приведен пример применения ПЛИС для построения гетерогенной супер-ЭВМ «СКИФ» ряда 4.

**Ключевые слова:** применение ПЛИС, реконфигурируемая вычислительная система, архитектурно-структурная организация супер-ЭВМ, базовый модуль на ПЛИС, открытая масштабируемая архитектура.

**Abstract.** Peculiarities of the architecture of supercomputers based on PLD are observed, the main principles of their construction are offered and peculiarities of architectural and structural organization of reconfigurable computing systems (RCS) are resulted, e.g.: in terms of basic modules based on PLD; RCS with an open scalable architecture (RCS OSA); with direct connections between computing modules and digit-by-digit processing of operands. The example of PLD application for development of SKIF supercomputers line (4th generation) is given.

**Keywords:** PLD application, reconfigurable computing system, architectural and structural organization of the supercomputer, basic modules based on PLD, an open scalable architecture.

### 1. Введение

Повышение производительности супер-ЭВМ – одна из постоянных проблем, решаемых при создании данного класса вычислительных машин. Это объясняется, главным образом, тем, что сложность задач пользователя постоянно увеличивается, соответственно возрастают и требования к скорости их решения, существенно опережая при этом возможности существующих средств обработки информации. В настоящий момент возможности повышения производительности системы за счет совершенствования интегральной технологии создания элементной базы практически исчерпаны, так как уменьшение размеров элементов на кристалле достигло своих пределов, когда уже появляются квантовые эффекты. Кроме того, с увеличением плотности размещения транзисторов на кристалле и сложности решаемых задач значительно возрастает сложность процессоров. При этом прирост производительности оказывается не пропорционален затрачиваемым аппаратным ресурсам и энергии. Заявленные производителями пиковые характеристики процессоров практически не достижимы без низкоуровневого программирования. Фактически апробирован также и

ряд архитектурно-структурных решений, значимость в повышении производительности которых стала не столь существенной. Например, применение иерархии КЭШ-памяти так же, как и увеличение её емкости, и увеличение числа одновременного выполнения команд уже слабо влияют на достижение поставленной цели, поскольку практически почти все ресурсы в этом направлении уже использованы. Часто не удовлетворяют пользователя и результаты объединения коммутационной средой однородных либо разнородных (гетерогенных) средств обработки информации в структурные организации, например, кластерные системы, GRID-системы и др., так как коммутационная среда и является причиной того, что большинство вычислительных задач выполняется с реальной производительностью не более 10–20% от заявляемой пиковой. Более того, производительность многопроцессорных вычислительных систем (МВС), ориентированных на традиционные методы организации параллельных вычислений, зачастую не превышает 10–15% от заявляемой пиковой вследствие необходимости реализации множества процедур межпроцессорного обмена, а также синхронизации последовательных процессов, выполняемых в процессорах системы. Для задач, требующих интенсивных, но нерегулярных межпроцессорных обменов и обращений к системе распределенной памяти, реальная производительность многопроцессорной вычислительной системы снижается до 1% [1]. Одной из основных причин такого явления есть несоответствие между “жесткой” архитектурой МВС и информационной структурой решаемых задач. Устранить этот недостаток можно путем построения реконфигурируемой вычислительной системы (РВС) с архитектурой, подстраиваемой под информационную структуру конкретной задачи. В соответствии с этим задача построения реконфигурируемых супер-ЭВМ является весьма актуальной.

## 2. Основные принципы построения РВС на ПЛИС

Реконфигурацию супер-ЭВМ под информационную структуру решаемой конкретной задачи (с различной степенью соответствия) можно выполнить следующим образом:

- путем выбора с помощью коммутационных средств необходимых ресурсов системы из заведомо введенных в систему избыточных программно-аппаратных средств [2];
- с помощью применения средств профессиональной ориентации (например, графических ускорителей), подключаемых через соответствующие интерфейсы к компьютерной системе [3];
- путем применения новых архитектурно-структурных решений, реализованных с использованием достижений твердотельной микроэлектроники для МВС типа “Processor-in-memory” (PIM-системы), “Memory-in-processor” (MIP-системы) и др., ориентированных на решение конкретной задачи при отсутствии ограничений на масштабирование [4, 5];
- путем использования программируемых логических интегральных схем (ПЛИС), обеспечивающих “гибкую” (программируемую) архитектуру МВС, подстраиваемую под информационную структуру конкретной задачи [6–13].

Применение ПЛИС в настоящее время является наиболее приемлемым для построения реконфигурируемых компонентов и супер-ЭВМ в целом, так как реализует наибольшее приближение архитектуры супер-ЭВМ к информационной структуре алгоритма решаемой задачи, обеспечивая при этом до 60–70 % реальной производительности системы по отношению к пиковой. Благодаря возможности ПЛИС интегрировать функцию обработки данных с помощью встроенного аппаратного и программно созданного процессорного блока, можно настраивать системы таким образом, чтобы удовлетворять требованиям специализированных приложений. Полученная таким образом система отличается от типового варианта реализации наличием высокоинтегрированной ПЛИС, которая включает в свою структуру встраиваемые микропроцессоры, оперативную память, а также блоки обработки данных, интерфейсные блоки и др.. При этом в качестве микропроцессорных блоков обработки используются (помимо имеющихся на кристалле стандартных блоков)

синтезируемые процессорные ядра (Soft Processor Cores), которые представляют собой СФ-блоки, разработанные и оптимизированные для систем, реализуемых на ПЛИС.

Основные принципы реконфигурации супер-ЭВМ при использовании ПЛИС в качестве базового элемента системы можно сформулировать следующим образом:

1. Принцип соответствия архитектурно-структурной организации системы, выполненной на ПЛИС, информационной структуре алгоритма, реализуемой на данной системе задачи пользователя.

Этот принцип фактически является основным при создании реконфигурируемых вычислительных систем (РВС) на ПЛИС, так как при более низких значениях параметров ПЛИС по сравнению с параметрами современных микропроцессоров (относительно низкая тактовая частота, меньшее число входов/выходов, меньший объем встроенной оперативной памяти и др.) такая система в состоянии обеспечить более чем на порядок повышение производительности, достигая при этом до 60–70 % реальной (рабочей) производительности системы по отношению к пиковой.

2. Системный принцип использования ПЛИС.

ПЛИС используется как свободный конфигурируемый ресурс, на базе которого реализуются необходимые устройства вычислительной системы. В качестве основных вычислительных элементов вместе с отдельными функционально законченными узлами и блоками, размещенными на кристалле ПЛИС, используются вычислительные структуры, созданные в поле логических блоков ПЛИС, а имеющийся универсальный микропроцессор (микропроцессоры), помимо выделенных для него арифметических операций, выполняет, как правило, и вспомогательные функции: загрузку конфигураций ПЛИС, обмен данными, управление, тестирование, загрузку исходных данных, визуализацию результатов и т.п. При этом реализация приложения должна быть легко масштабируемой и обеспечивать соответствующий прирост производительности вычислительной системы при увеличении аппаратных ресурсов. Кроме того, должны быть обеспечены совместимость и переносимость проектов между реконфигурируемыми вычислительными структурами разных архитектур.

3. Принцип архитектурно-структурной организации системы на ПЛИС для поддержки прогрессивных информационных технологий.

Используемые и сформированные внутри ПЛИС средства обработки информации должны реализовывать параллельно-конвейерную схему обработки потока данных/инструкций с минимальной временной задержкой между ступенями конвейера. При этом обработка потока данных/инструкций должна происходить с интенсивностью, не меньшей интенсивности их поступления.

В качестве перспективного направления создания информационной структуры РВС целесообразно использовать структуру базового модуля РВС с открытой архитектурой (РВС ОМА), с ортогональной системой связей и кольцевым информационным каналом, представляющим собой множество независимых высокочастотных каналов.

4. Принцип применения структурно-процедурного способа организации вычислительного процесса.

При реализации структурно-процедурного способа информационный граф задачи представляется в виде кортежа изоморфных базовых подграфов, которые являются информационно независимыми или непосредственно зависящими друг от друга. Кортеж информационных подграфов преобразуется в специальную вычислительную конструкцию – кадр, которому соответствует подграф задачи, реализованный аппаратно и через который следует поток операндов. При этом каждая группа операндов (результатов) соответствует входным (выходным) вершинам определенного подграфа кортежа. Смена кадров в РВС осуществляется процедурно по единой для всей системы программе. Механизм последова-

тельного обхода подграфов информационного графа задачи кадрами принято называть [1] структурно-процедурной организацией вычислений.

5. Принцип пакетной передачи данных внутри ПЛИС и с внешними по отношению к ПЛИС конечными устройствами.

Передача данных выполняется пакетами, поддерживая различные протоколы взаимодействия конечных устройств (отложенные транзакции – posted, немедленные транзакции – nonposted, атомарные операции – atomic operation, транзакции барьерной синхронизацией – flush/fence transaction и др.). При этом обеспечивается вероятность переполнения буферов входящих/выходящих пакетов близкой к нулю.

6. Принцип многоуровневого программного обеспечения систем, реализованных на ПЛИС.

Программирование РВС можно условно разделить на две составляющие. Первая – структурная составляющая – предполагает создание в базовой архитектуре РВС специализированной вычислительной структуры, адекватно соответствующей решаемой задаче. Вторая – процедурная составляющая – организует вычислительный процесс в созданной структуре.

В общем случае при реконфигурации на ПЛИС выделяют 4 основных уровня программирования реконфигурируемых вычислительных систем в прикладной масштабируемой программе [6]: уровень использования функционально законченных фрагментов; уровень программирования унифицированного макрообъекта для прикладной масштабируемой программы; уровень синтеза и программирования макрообъектов в прикладной масштабируемой программе; уровень трансляции прикладной масштабируемой программы с языка высокого уровня в логические ячейки ПЛИС и связи между ними. При этом создается (выбирается) программное обеспечение для управления аппаратурой и предоставления интерфейса программирования приложений (API) для разработчиков прикладных программ; обеспечивается программирование всех составляющих масштабируемых прикладных программ на языке высокого уровня; обеспечивается реконфигурация прикладных масштабируемых программ без участия высококвалифицированного специалиста-схемотехника, что при отсутствии такового в некотором роде создает трудноразрешимую проблему.

Программирование РВС на первом уровне выполняется с помощью вызова внешних библиотечных функций из программ на традиционных высокоуровневых языках программирования [6]. Вызываемая функция осуществляет запуск потоков данных, следующих через вычислительную систему. При этом структурная, потоковая и процедурная составляющие прикладной масштабируемой программы не меняются и должны быть уже загружены в РВС.

Программирование на уровне унифицированных макрообъектов (представляющих собой совокупность вычислительных устройств, выполняющих определенную группу команд и соединенных между собой коммутационной системой) позволяет программисту задавать коммутацию как внутри макрообъектов, так и макрообъектов между собой, что, в свою очередь, позволяет перестраивать структуру РВС в процессе решения задачи и обеспечивает пользователя более гибкими средствами разработки прикладных программ.

Программные средства для поддержки третьего уровня программирования обеспечивают возможность адаптации программных компонентов средств разработки для РВС при переходе на новые топологии. Поддержка вводимых расширений всеми средствами разработки прикладных программ на всех уровнях позволяет создавать эффективные прикладные программы для РВС при решении задач различных предметных областей, обеспечивая удобство программирования и сокращая время перетрансляции прикладного решения.

### 3. Особенности организации архитектур РВС, выполненных на основе базовых модулей, реализованных на ПЛИС

В основу построения семейства РВС положены архитектурные принципы модульно-наращиваемых многопроцессорных вычислительных систем с программируемой архитектурой и структурно-процедурной обработкой информации.

В простейшем случае структурно-процедурный вычислительный процесс в РВС организуется следующим образом. В аппаратном ресурсе РВС реализуется один из базовых подграфов сегментированного информационного графа задачи, и на его входы из блоков памяти подаются входные данные этого фрагмента. Процесс вычисления реализуется вычислительной структурой, соответствующей данному подграфу, и на выходе формируется результат вычислений, который запоминается в соответствующих блоках памяти. Далее в аппаратном ресурсе РВС реализуется следующий подграф сегментированного информационного графа задачи, и процесс повторяется [1, 8].

Укрупненная структурная схема РВС, реализующей этот процесс, приведена на рис. 1 [8], ориентирована на аппаратную реализацию всех операций с помощью макропроцессоров (МАП), предписанных вершинами информационного графа задачи всех каналов передачи данных между вершинами соответствующих дугам графа, и всех информационных каналов, соответствующих входным и выходным вершинам. Каждый макропроцессор содержит набор элементарных процессоров (ЭП), объединенных в единый вычислительный ресурс с помощью локального пространственного коммутатора ( $K_1$ ) (рис. 1а), который соединяет по полному графу все информационные входы и выходы ЭП, внешние информационные входы  $X$ , внешние информационные выходы  $Z$  макропроцессора. Настройка макропроцессора на макрооперацию производится с помощью блока макроопераций (БМ).

Обобщенная схема РВС приведена на рис. 1б. В состав РВС входят: распределенная память (РП), обеспечивающая возможности параллельной выдачи массивов входных данных на входы макропроцессоров и записи результатов вычислений с их выходов; множество МАП, состоящих из ЭП, соединяемых между собой с помощью локального коммутатора  $K_1$ ; множество каналов распределенной памяти, в каждый из которых входит контроллер распределенной памяти (КРП). Множество входных и выходных дуг информационного графа реализуются системным коммутатором  $K_2$ , обеспечивающим подключение каналов распределенной памяти к тем или иным входам и выходам МАП. При этом массивы входных данных отображаются в соответствующие каналы распределенной памяти, процедура обращения к которым реализуется с помощью КРП [8].

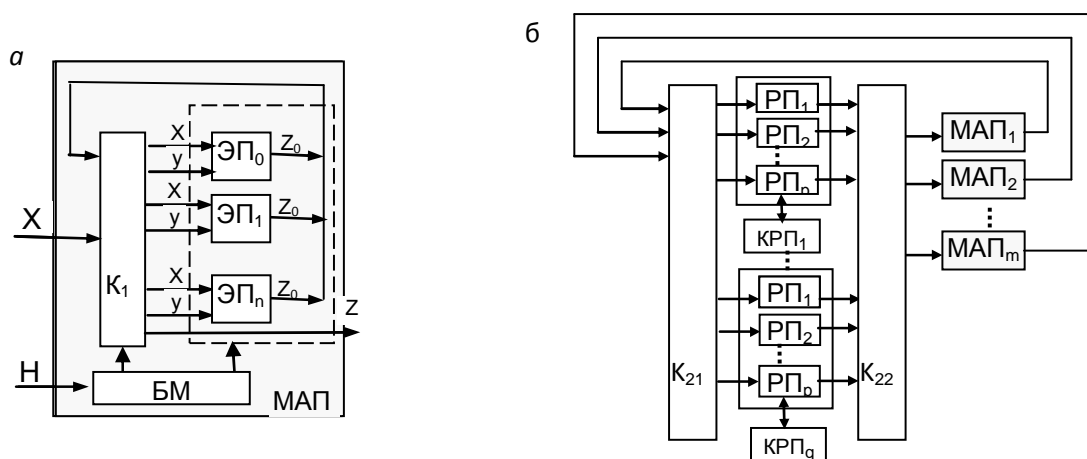


Рис. 1. Структурная схема макропроцессора (рис. 1а) и обобщенная схема РВС (рис. 1б)

Особенность архитектуры РВС заключается в том, что в процессе конструирования она не формируется окончательно, а остается в определенном смысле незавершенной и открытой. Окончательное программирование архитектуры РВС включает создание функциональных узлов (элементарных процессоров) для выполнения вычислений, настройку прямых информационных каналов в коммутаторах  $K_1$  и  $K_2$  между элементарными процессорами, настройку блоков распределенной памяти на реализацию процедур чтения и записи информационных массивов. Аппаратно-программные средства РВС позволяют синтезировать произвольные соединения между компонентами системы, создавая необходимые вычислительные структуры. Таким образом, совокупность вычислительных структур, созданных в рамках базовой архитектуры РВС, образуют виртуальный проблемно-ориентированный вычислитель, структура которого адекватна информационному графу (подграфу) решаемой задачи.

Применение ПЛИС большой степени интеграции дало возможность перейти к созданию в РВС больших решающих полей, содержащих множество однотипных микросхем. При этом эквивалентное решающее поле является как бы аналогом сверхбольшой ПЛИС, объединяющей ресурсы всех микросхем. Архитектурные особенности базового модуля и модульно наращиваемой РВС на его основе показаны на рис. 2а и 2б соответственно [8].

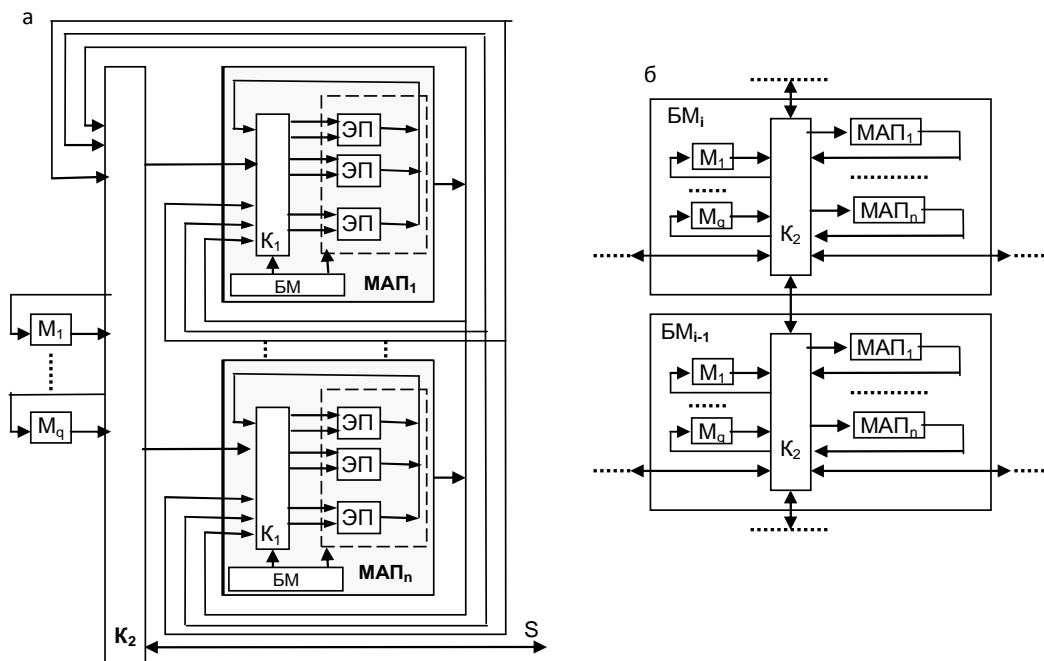


Рис. 2. Укрупненная структура базового модуля (рис. 2а) и обобщенная структура модульно-наращиваемой РВС (рис. 2б)

В базовом модуле (рис. 2а) используется двухуровневая коммутационная система. Первый уровень устанавливает связи между  $ЭП$  внутри  $МАП$ , а второй уровень – связи между  $МАП$  внутри базового модуля. При этом контроллеры распределенной памяти (КРП) и сами блоки распределенной памяти (РП) выполняются на типовых микросхемах ОЗУ необходимого объема и быстродействия. Контроллер базового модуля (КБМ) может быть реализован на ПЛИС и выполняет функции управления и контроля всех систем базового модуля. На базовом модуле располагаются также и вспомогательные подсистемы: синхронизации, электропитания и охлаждения. Данные между дальними микросхемами передаются по транзитным каналам через промежуточные микросхемы, используя систему ортогональных связей.

Контроллеры памяти работают с фрагментами параллельной программы, загруженными в их блоки памяти, исполняя которые, контроллеры участвуют в настройке элементарных процессоров на выполнение необходимых операций и создают с помощью коммутаторов  $K1$  и  $K2$  необходимые каналы связи между ними, тем самым реализуя в пределах базового модуля мультиконвейерную вычислительную структуру, соответствующую базовому подграфу задачи. Нарращивание производительности РВС достигается путем объединения нескольких базовых модулей в единый вычислительный ресурс (рис. 2 б) [8].

#### 4. Состав и параметры семейства РВС на основе ПЛИС

Концепция построения РВС на основе ПЛИС позволила создать на единых архитектурных принципах целый ряд высокопроизводительных систем различных архитектур и конфигураций. На рис. 3 отображены названия некоторых из них, при этом РВС, в которых основные узлы и блоки выполнены на ПЛИС, рассмотрены несколько подробнее.

НИИ многопроцессорных вычислительных систем Южного федерального университета (НИИ МВС ЮФУ, Россия) выполнил ряд работ по созданию реконфигурируемых вычислительных систем, в которых множество ПЛИС образует единый вычислительный ресурс, аппаратно реализующий трудоемкие фрагменты решения задачи [1, 6–10].

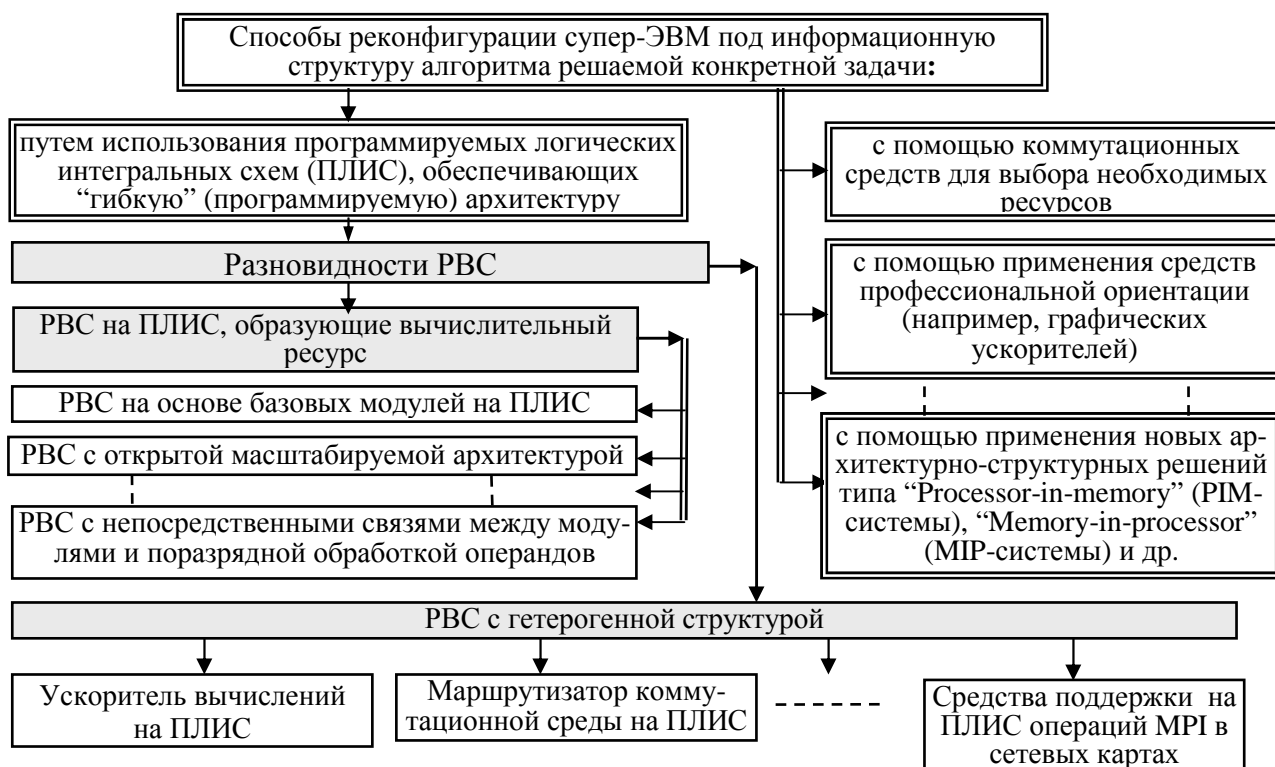


Рис. 3. Способы реконфигурации и разновидности реконфигурируемых вычислительных систем

Среди них следует выделить семейство программно-совместимых РВС производительностью от 0,025 Тфлопс до 6 Тфлопс, которое включает:

- РВС-5 – высокопроизводительную систему производительностью 6 Тфлопс;
- РВС-1Р и РВС-1К – системы производительностью более 1 Тфлопс;
- РВС-0.2-РС – рабочую станцию производительностью 300 ГФлопс;
- РУПК-50 и РУПК-25 – ускорители персональных компьютеров производительностью 50 и 25 ГФлопс. Пример основных параметров отдельных РВС на ПЛИС приведен в табл. 1.

Таблица 1. Основные параметры РВС высокой производительности, выполненные на ПЛИС

Наименование параметра	Рабочие станции		Ускорители ПЭВМ	
	РВС-02-РС	РВС-1Р	РУПК-25	РУПК-50
Количество БМ, шт.	4	16	1-БМ “16S3-25”	1-БМ “16V5-50”
Количество ПЛИС, шт.	64	256	16	16
Количество ЭП, шт.	1024	4096	128	256
Количество эквив. вентилях, шт.	$0,7 \times 10^9$	$2,8 \times 10^9$	$64 \times 10^6$	$176 \times 10^6$
Производительность, Гфлопс	300	1000	25	50
Объем оперативной памяти, Гбайт	5,6	22	1,5	1,5
Потребляемая мощность, Вт (не более)	1200	4800	200	300

Базовый модуль 16S3-25 [7] ускорителя РУПК-25 конструктивно состоит из печатной платы базового модуля 16S3-25, которая содержит 16 ПЛИС серии Spartan 3 фирмы Xilinx емкостью  $4 \cdot 10^6$  эквивалентных вентилях каждый. Состав базового модуля позволяет реализовать 128 устройств с плавающей запятой с 64-разрядным форматом данных или 256 устройств с плавающей запятой с 32-разрядным форматом данных, которые связаны между собой и каналами распределенной памяти пространственной коммутационной системой. Решающее поле базового модуля 16V5-50 ускорителя РУПК-50 выполнено на 16-ти ПЛИС Virtex 5 XC5VLX110-1FFG1153 фирмы Xilinx, контроллер базового модуля – на ПЛИС XC5VLX50T-1FFG1136. Распределенная память реализована на 24 микросхемах SDRAM типа DDR2 и имеет общий объем 1,5 Гбайта. Производительность базового модуля 16V5-50 составляет 50 Гфлопс.

Старшие представители семейства создаются на принципах модульной наращиваемости на основе базовых модулей и соответствующих вычислительных блоков. При этом они, как правило, обладают почти линейным ростом реальной производительности в зависимости от увеличения аппаратного ресурса.

Основу РВС-1Р составляют вычислительная стойка СТ-1Р, в которую монтируются четыре вычислительных блока РВС-0.2-ВБ, ПЭВМ, коммутатор Ethernet, система питания и система охлаждения. Структурная схема СТ-1Р показана на рис. 4 [1]. Реконфигурируемая вычислительная система РВС-5 является одним из представителей семейства высокопроизводительных многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой на основе реконфигурируемой элементной базы. РВС-5 содержит [9] 5 вычислительных стоек РВС-1, 20 вычислительных блоков РВС-0.2-ВБ, 80 базовых модулей 16V5-75, 25600 элементарных процессоров и работает на тактовой частоте 330 МГц, обрабатывая 64-разрядные данные с производительностью более 6000 Гфлопс. При этом обмен между блоками осуществляется с частотой 640 МГц. Система имеет внешние интерфейсы типа LVDS и Gigabit Ethernet.

Результаты исследований показали, что применение ПЛИС семейства Virtex-6 в качестве элементной базы для построения вычислительных модулей при сохранении стоимости вычислительного модуля позволяет увеличить производительность в 1,5-2 раза по сравнению с аналогичным решением на основе ПЛИС семейства Virtex-5, а применение ПЛИС семейства Virtex-7 улучшает характеристики примерно в 1,7 раза по сравнению с изделиями на основе Virtex-6 [6, 9]. Этот факт позволяет рассматривать создание вычислительных модулей на современных ПЛИС последних выпусков в качестве наиболее перспективного направления для построения РВС, так как обеспечивает им конкурентное преимущество по большинству технико-экономических показателей.



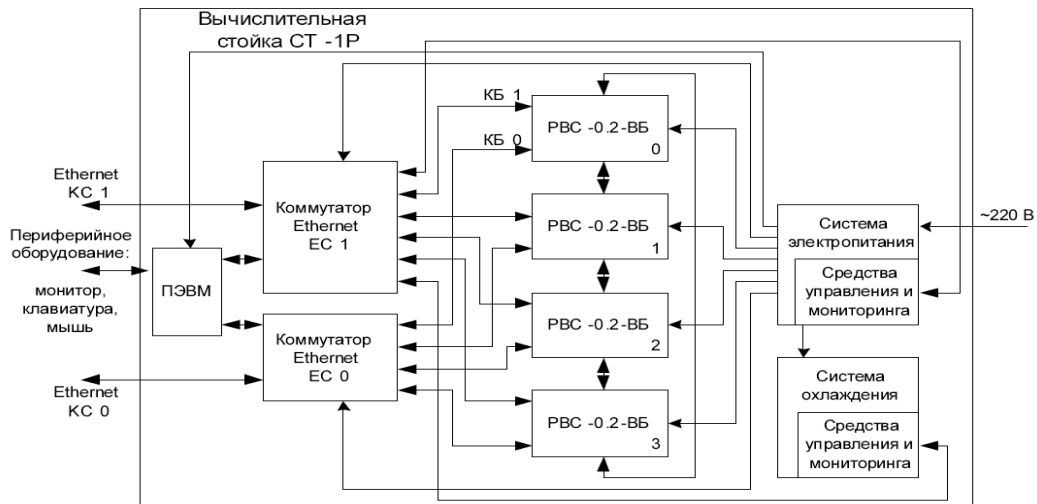


Рис. 4. Увеличенная структурная схема СТ-1Р

В соответствии с этим кратко рассмотрим состав и параметры РВС-7, построенной на основе ПЛИС семейства Xilinx Virtex-7. Для его создания используется перспективный вычислительный модуль 24V7-750 [6, 10] на ПЛИС Virtex-7. Этот модуль содержит 4 платы вычислительного устройства (ПВМ) 6V7-180 и обеспечивает РВС-7 пиковую производительность до  $10^{15}$  операций с фиксированной запятой в секунду в одностоечном конструктиве. Для связи с другими ВМ используются 12 каналов интерфейса LVDS на частоте 800 МГц по 25 дифференциальных пар каждый (разъёмы типа SS4). Объем распределенной памяти ВМ 24V7-750 составляет 12 Гбайт.

Для программирования РВС-7 используется комплекс программного обеспечения, одним из основных инструментов которого является специально разработанный язык программирования высокого уровня COLAMO [10], предназначенный для описания реализации параллельного алгоритма. Основой способов является программирование макрообъектов, которое включает в себя загрузку во все используемые в РВС макрообъекты управляющего пакета с целью их настройки на параметры решаемой задачи. Разработанный на основе ПЛИС Xilinx Virtex-7 вычислительный модуль 24V7-750 и созданная на его основе РВС-7 вместе с необходимыми программными средствами и указанными аппаратными ресурсами обеспечивают создание проблемно-ориентированных и специализированных вычислительных структур с высокой эффективностью вычислений при решении широкого класса задач.

## 5. Концепция построения РВС с открытой масштабируемой архитектурой (РВС ОМА)

По мере роста интеграции ПЛИС появилась возможность существенно увеличить количество элементарных процессоров на ПЛИС, однако соответственно увеличить количество блоков распределенной оперативной памяти и соответствующих каналов памяти, оставаясь в рамках ортогональной архитектуры, оказалось невозможным. Возникла необходимость поиска новой архитектуры РВС и оригинальных конструктивно-технологических решений для создания базовых модулей, которые бы исключили отрицательное влияние этих факторов. В результате появилась концепция построения РВС с открытой масштабируемой архитектурой (РВС ОМА). При этом в качестве основы для ОМА предлагается использовать информационную структуру базового модуля, упрощенная схема которого представлена на рис. 5 [11].

Базовый модуль РВС ОМА содержит множество ПЛИС, связанных с помощью двух

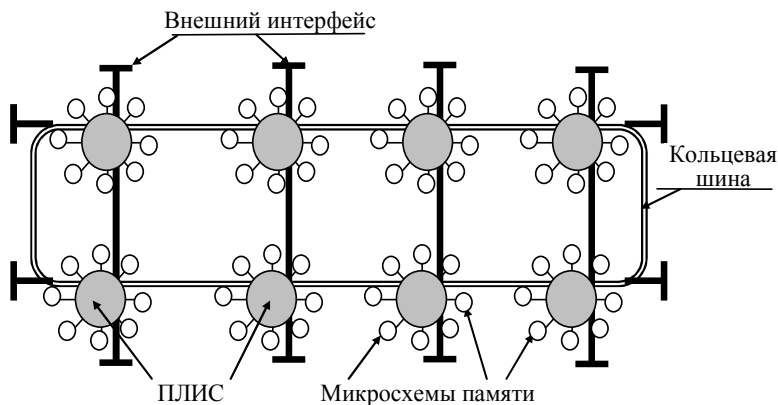


Рис. 5. Условное обозначение информационной структуры РВС с открытой архитектурой

типов связей. К внешним разъемам ведут LVDS-каналы, обозначенные вертикальными линиями.

Такие же каналы образуют соединения между линейками ПЛИС. На данных линиях связи обеспечивается утроение тактовой частоты информационных обменов по сравнению с частотой обработки информации. Однако эти связи являются вспомогательными, так как основную ин-

формационную нагрузку обеспечивает кольцевой информационный канал, представляющий собой множество независимых высокочастотных каналов. Второе ограничение относительно количества выводов микросхемы ПЛИС может быть в принципе решено вследствие многократного увеличения частоты передачи данных в кольцевом информационном канале, обеспечивающем возможность увеличения количества виртуальных выводов ПЛИС по сравнению с ортогональной архитектурой примерно на порядок. При этом пользователь должен воспринимать вычислительную систему как совокупный реконфигурируемый ресурс, на который он посредством трансляторов и синтезаторов будет отображать вычислительную структуру.

Таким образом, РВС ОМА обладает высоким потенциалом и может использоваться в качестве базовой архитектуры для эффективного решения задач любой проблемной области.

## 6. Особенности построения РВС с непосредственными связями между вычислительными модулями и поразрядной обработкой операндов [12].

Применение вычислительных модулей, обрабатывающих операнды поразрядно, позволяет совмещать выполнение зависимых по данным операций и сократить число связей между модулями. Вычислительный модуль (ВМ) с поразрядной передачей данных использует существенно меньше ресурсов ПЛИС. При этом в несколько раз экономятся как внутренние ресурсы ПЛИС, так и ее выводы. Это дает возможность реализовать на той же микросхеме ряд других устройств, и тем самым обеспечивается построение системы на одной ПЛИС, что, естественно, повышает ее надежность, уменьшает энергопотребление и габариты, а также обеспечивает потенциальную возможность повысить частоту тактирования. Это, в свою очередь, ускоряет обработку информации. При обработке данных, кроме сокращения числа связей, появляется возможность выполнять зависимые по данным операции в режиме частичного совмещения. При таком режиме вычислений выполнение следующей операции будет начинаться не после завершения выполнения предыдущей операции, а сразу же после получения первого разряда результата этой операции. Режим работы таких ВМ называют неавтономным, так как для выполнения последовательности операций необходимо соединить несколько ВМ, которые совместно выполняют цепочку операций, обмениваясь информацией в процессе работы. Такие ВМ по структуре ближе к параллельным, а не последовательным устройствам, что определило их название «квазипараллельные».

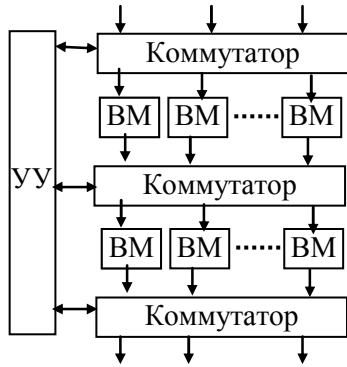


Рис. 6. Структура системы с перестраиваемыми связями между ВМ

В реконфигурируемых ВС (рис. 6) необходимое соединение между ВМ обеспечивает коммутационная среда, выполненная на ПЛИС, которая настраивается в соответствии с графом потока данных (ГПД). Разряды результата выдаются со старших разрядов, причем первый разряд формируется с задержкой на  $p$  шагов. Следовательно, число шагов  $N$ , необходимое для получения  $n$  старших разрядов окончательного результата при выполнении цепочки из  $K$  операций, составляет [12]

$$N = n - 1 + \sum_j^K (p_j + 1),$$

где  $j$  – индекс операции, лежащей на критическом пути в графе алгоритма, а  $p_j$  – задержка формирования результата при выполнении  $j$ -й операции.

Сокращение занимаемой площади на кристалле ПЛИС создает возможность построения системы на одной ПЛИС, что обеспечивает повышение надежности РВС, уменьшение энергопотребления и габаритов, а также дает потенциальную возможность повысить частоту тактирования, что, в свою очередь, ускоряет обработку информации.

## 7. Повышение производительности гетерогенных (гибридных) структур супер-ЭВМ с помощью ПЛИС

В настоящее время наблюдается переход на гибридные, гетерогенные структуры суперкомпьютеров, узлы которых состоят как из стандартных процессоров, так и из ПЛИС. В общем случае подходы к повышению производительности гетерогенных структур супер-ЭВМ с помощью ПЛИС можно подразделить на два класса [13]:

- 1) реализация в ПЛИС общесистемных операций нижнего уровня для конкретной модели вычислений; при этом ПЛИС берет на себя обслуживание прикладных программ, разработанных в данной системе программирования, а реализация непосредственно прикладных программ выполняется на стандартном процессоре;
- 2) реализация в ПЛИС критичных по эффективности фрагментов прикладных программ или их целиком. При этом для каждой прикладной программы загружается новая схема в ПЛИС.

Ни более перспективными являются архитектуры, в которых ПЛИС используется в качестве маршрутизатора коммутационной среды; средства поддержки базовых операций, а также в качестве ускорителя вычислений. Среди работ по созданию различных моделей вычислений и их аппаратной поддержке можно отметить следующие:

1. Гибридные суперкомпьютеры, содержащие в каждом узле ускоритель вычислений с нестандартной архитектурой, реализованный на ПЛИС совместно с маршрутизатором коммутационной среды.
2. Средства аппаратно-программной поддержки в сетевых картах операций MPI для перспективных моделей вычислений на базе ПЛИС.
3. Реконфигурируемые многопроцессорные вычислительные системы, разработанные НИИ МВС ЮФУ (Россия), в которых реализован структурно-процедурный способ организации мультиконвейерных вычислений, сущность которого состоит в том, что приложение автоматически разбивается транслятором на компоненты, которые исполняются на специальной системе, состоящей из ПЛИС и процессоров. Это потребовало создания качественно новой методики программирования.

4. Другие работы, отмеченные в [13].

В суперкомпьютерах «СКИФ» ряда 4 задача разработки современных суперкомпьютеров с использованием передовых технологий, а также проблема сочетания новых и классических архитектур решается ИПМ РАН путем применения в вычислительных узлах процессоров Intel® Xeon® 5570 (Nehalem) и соответствующего им чипсета, а также ПЛИС Altera EP4SGX230 и других микросхем с программируемой структурой. Технические характеристики вычислительного узла приведены в табл. 2 [13].

Таблица 2. Технические характеристики вычислительного узла «СКИФ» ряда 4 и основных коммуникационных сетей

Наименование параметра	Значение параметра
Тип процессора	4-ядерный Intel® Xeon® CPU X5570 @ 2.93GHz
Количество процессоров	2
Пиковая производительность	93.76 Гфлоп/с (на 2 процессора узла)
Тип сопроцессора	ПЛИС Altera EP4SGX230
Топология системной сети	Трехмерный тор (3D-тор)
Максимальная скорость передачи данных системной сетью (3D-тор)	60 Гбит/с/узел (суммарная на 6 линках)
Задержка передачи сообщений на уровне прикладных программ (MPI)	1–2 мкс
Тип вспомогательной сети	Mellanox ConnectX InfiniBand HCA QDR
Максимальная скорость передачи данных вспомогательной сетью InfiniBand	40 Гбит/с/узел
Дисковая память	Накопитель SSD 80 Гбайт
Оперативная память	12 Гбайт

Суперкомпьютеры «СКИФ» ряда 4 содержат две основные коммуникационные сети:

1. Вспомогательная сеть на базе серийного сетевого оборудования Mellanox ConnectX Infiniband HCA QDR. Эта сеть обеспечит стандартные режимы работы, доступ к параллельной файловой системе суперкомпьютера, функционирование средств управления суперкомпьютером. Максимальная пропускная способность этой сети – 40 Гбит/с/узел.

2. Системная сеть на базе ПЛИС с линками, объединенными в сеть с топологией 3D-тор. Эта сеть поддерживает разработку новых эффективных методов маршрутизации, реализацию операций, связанных с обменом, новых моделей параллельных вычислений и т.п. Максимальная пропускная способность этой сети в полтора раза больше вспомогательной – 60 Гбит/с/узел.

## 8. Выводы

Применение ПЛИС для создания реконфигурируемых вычислительных систем (РВС) является принципиально новым направлением в построении супер-ЭВМ, в которых в качестве основного вычислительного элемента используются не универсальные микропроцессоры, а ПЛИС-микросхемы. Это дает возможность пользователям создавать в базовой архитектуре РВС виртуальные специализированные вычислители, структура которых адекватна структуре решаемой задачи. Поэтому, несмотря на более низкую частоту в ПЛИС, чем у современных микропроцессоров, и меньшее число эквивалентных вентилях, ПЛИС обеспечивает ускорение вычислительного процесса по сравнению с микропроцессором от 2 до 10 раз. При этом применение структурно-процедурного принципа организации вычисле-

ний позволяет довести реальную производительность до 60% от пиковой производительности на широком классе задач и обеспечивает практически линейный рост производительности при наращивании аппаратного ресурса.

Применение информационной структуры базового модуля РВС с открытой архитектурой (РВС ОМА), с ортогональной системой связей и кольцевым информационным каналом, представляющим собой множество независимых высокочастотных каналов, позволяет существенно упростить печатную плату и улучшить ее частотные характеристики, при этом обеспечивая 10-кратное увеличение числа виртуальных выводов ПЛИС.

Разработанная и созданная на основе микросхемы ПЛИС Xilinx Virtex-7 система РВС-7 с необходимыми программными средствами показала, что на настоящее время возможно создание высокопроизводительной РВС, обеспечивающей высокую эффективность вычислений на задачах различных классов.

В основу построения семейства РВС положены архитектурные принципы модульно-наращиваемых многопроцессорных вычислительных систем с программируемой архитектурой и структурно-процедурной обработкой информации.

Принятая в «СКИФ» ряда 4 гибридная архитектура обеспечивает уникальную возможность поддержки в ПЛИС как новых моделей параллельных вычислений, так и старых систем программирования, особенно в части эффективной коммуникации и доступа к распределенной памяти.

## СПИСОК ЛИТЕРАТУРЫ

1. Левин И.И. Семейство высокопроизводительных реконфигурируемых вычислительных систем / И.И. Левин // Вестник Томского государственного университета: Управление, вычислительная техника и информатика. – 2008. – № 2 (3). – С. 77 – 93.
2. Пат. № 6259 Украина, МПК G06F13/00, G06F12/00. Система пам'яті з інтеграцією функцій зберігання та обробки інформації на одному кристалі / Сергієнко І.В., Кривонос Ю.Г., Палагін О.В., Коваль В.М., Яковлев Ю.С., Тихонов Б.М.; Інститут кібернетики Імені В.М. Глушкова НАН України; опубл. 15.04.05, Бюл. № 4. – 24 с.
3. Яковлев Ю.С. О выборе графических ускорителей для компьютерных систем / Ю.С. Яковлев // Наукові праці ДОННТУ. – (Серія “Інформатика, кібернетика та обчислювальна техніка”). – 2013. – №2 (18). – С. 61 – 71.
4. Пат. на винахід 99164 Україна, МПК G06F 15/16, G06F 13/42. Інтелектуальна розподілена система пам'яті з кільцевою шиною / Палагін О.В., Яковлев Ю.С., Тихонов Б.М., Єлісеєва О. В.; Інститут кібернетики ім. В.М. Глушкова НАН України; заявл. 16.07.10; опубл. 25.07.12, Бюл. № 14. – 21с.
5. Яковлев Ю.С. Однокристалльные компьютерные системы высокой производительности. Особенности архитектурно-структурной организации и внутренних процессов / Яковлев Ю.С. – Винница: ВНТУ, 2009. – 294 с.
6. Перспективные реконфигурируемые вычислительные системы на основе ПЛИС Virtex-6 и Virtex-7 [Электронный ресурс] / И.И. Левин, И.А. Каляев, А.И. Дордопуло [и др.]. – Режим доступа: <http://agora.guru.ru/abrau2012/pdf/338.pdf>.
7. Каляев И.А. Реконфигурируемые вычислительные системы с высокой реальной производительностью [Электронный ресурс] / И.А. Каляев, И.И. Левин. – Режим доступа: <http://www.ict.edu.ru/yconf/files/11875.pdf>.
8. Каляев.И.А. Архитектура семейства реконфигурируемых вычислительных систем на основе ПЛИС / И.А. Каляев, И.И. Левин, Е.А. Семерников // Штучний інтелект. – 2008. – № 3. – С. 663 – 673.
9. РВС-5. – Режим доступа: <http://superevm.ru/index.php?page=vychislitel'naya-sistema-rvs-5-2009>.
10. Реконфигурируемая вычислительная система на основе ПЛИС Vrtex-7 / И.А. Каляев, И.И. Левин, А.И. Дордопуло [и др.]. – Режим доступа: <http://2012.nscf.ru/Tesis/Levin.pdf>.
11. Левин И.И. Реконфигурируемые вычислительные системы с открытой масштабируемой архитектурой / И.И. Левин // Доклады Пятой междунар. конф. «Параллельные вычисления задачи управления». – Москва, 2010. – С. 83 – 95.

12. Жабин В.И. Эффективность реализации потоковых вычислений в системах с непосредственными связями на ПЛИС / В.И. Жабин, В.В. Жабина, М.А. Безгинский // Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка. – 2009. – № 55. – С. 149 – 156.
13. Возможности суперкомпьютеров «СКИФ» ряда 4 по аппаратной поддержке в ПЛИС различных моделей параллельных вычислений [Электронный ресурс] / С.М. Абрамов, С.А. Дбар, А.В. Климов [и др.]. – Режим доступа: <http://skif.pereslavl.ru/psi-info/rcms/rcms-publications/2010-rus/divn-abram-i-dr-2010.pdf>.

*Стаття надійшла до редакції 03.12.2013*