

РЕКОНФИГУРИРУЕМЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ ДЛЯ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО ТРУДОЕМКИХ ЗАДАЧ

Н.Н. Дмитренко, И.А. Каляев, И.И. Левин, Е.А. Семерников

ВВЕДЕНИЕ

Наиболее распространенные в мире кластерные МВС, предназначенные для решения больших задач, имеют высокие значения пиковой производительности. В то же время пользователи отмечают принципиальные недостатки кластерных МВС, которые связаны с относительно низкой скоростью процедур межпроцессорного обмена, ограниченной пропускной способностью сети передачи данных, необходимостью синхронизации множества взаимосвязанных последовательных процессов, каждый из которых выполняется на отдельном процессоре, и т.д. Все это приводит к тому, что высокую реальную производительность кластерные суперЭВМ демонстрируют, в основном, только при решении класса слабосвязанных задач, не требующих большого числа информационных обменов, в то время как при решении задач других классов их реальная производительность существенно снижается и не превышает 5-10% от декларируемой пиковой производительности системы. Это является следствием неадекватности данной конкретной архитектуры суперкомпьютера информационной структуре решаемой задачи [1, 2, 3].

Недостатки кластерных систем, связанные с их неизменной «жесткой» архитектурой, позволяет преодолеть концепция создания МВС с «гибкой», динамически перестраиваемой (программируемой) архитектурой на основе реконфигурируемой элементной базы или иначе – реконфигурируемые вычислительные системы (РВС) [3, 4]. Концепция создания РВС разработана и развивается в Научно-исследовательском институте многопроцессорных вычислительных систем имени академика А.В. Каляева Южного федерального университета (НИИ МВС ЮФУ) уже более 20 лет. В отличие от многопроцессорных вычислительных систем с «жесткой» архитектурой архитектура РВС может динамически изменяться в процессе их функционирования. В результате у пользователя появляется возможность адаптации архитектуры вычислительной системы под структуру решаемой им задачи. Иными словами, пользователь может, оставаясь в рамках базовой архитектуры системы, создавать проблемно-ориентированные вычислители, структура которых соответствует структуре решаемой задачи. Проведенные исследования и многочисленные практические разработки, выполненные в НИИ МВС ЮФУ, показали [5,6], что реализация данной концепции обеспечивает высокую реальную производительность РВС, близкую к пиковой, на широком классе задач, в том числе при решении «сильносвязанных» задач.

СОСТАВ И ХАРАКТЕРИСТИКИ СЕМЕЙСТВА РВС ДЛЯ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО ТРУДОЕМКИХ ЗАДАЧ

В настоящее время по заданию Федерального агентства по науке и инновациям ведется выполнение Государственного контракта № 02.524.12.4002 по теме «Создание семейства высокопроизводительных многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой на основе реконфигурируемой элементной базы и их математического обеспечения для решения вычислительно трудоемких задач» в рамках Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2012 годы».

Целью разработки является создание на единых архитектурных принципах семейства программно-совместимых реконфигурируемых высокопроизводительных вычислительных систем производительностью от 0,025 Тфлопс до 6 Тфлопс.

В рамках государственного контракта будет создано семейство РВС:

1. РВС-5 – высокопроизводительная система производительностью 6 Тфлопс;
2. РВС-1Р и РВС-1К – системы производительностью более 1 Тфлопс;
3. РВС-0.2-РС – рабочая станция производительностью 300 ГФлопс;
4. РУПК-50 и РУПК-25 – ускорители персональных компьютеров производительностью 50 и 25 ГФлопс.

Старшие представители семейства создаются на принципах модульной наращиваемости и будут обладать почти линейным ростом реальной производительности в зависимости от увеличения аппаратного ресурса [7].

Для создания семейства РВС, системного и прикладного программного обеспечения его представителей, информационной и программной инфраструктуры для обучения и поддержки потенциальных пользователей семейства образован консорциум из ряда научных и научно-производственных организаций, в который вошли: НИИ МВС ЮФУ (г. Таганрог) – головной исполнитель проекта; ФГУП «НИИ «Квант» (г. Москва); «Южный научный

центр РАН» (г. Ростов-на-Дону); «Специальная астрофизическая обсерватория РАН» (пос. Нижний Архыз); НИЦ «СуперЭВМ и нейрокомпьютеров» (г. Таганрог); НИВЦ МГУ (г. Москва).

Реальная производительность всех представителей семейства PBC планируется не ниже 50% от указанной пиковой производительности на широком классе вычислительных задач. При этом они будут обладать существенно меньшими габаритами, энергопотреблением и стоимостью по сравнению с суперЭВМ кластерного типа аналогичной производительности.

БАЗОВЫЕ МОДУЛИ СЕМЕЙСТВА PBC

Базовые модули, из которых строятся PBC, в полной мере сохраняют все концептуальные особенности архитектуры реконфигурируемых систем. Семейство PBC строится всего из трех типов базовых модулей: 16V5-75, 16V5-50 и 16S3-25. Базовый модуль 16V5-75, как самый производительный, используется в моделях PBC-5, PBC-1P и PBC-0.2-PC. Базовые модули 16V5-50 и 16S3-25 входят в состав ускорителей персональных компьютеров РУПК-50 и РУПК-25. Общие принципы построения базовых модулей одинаковы и поэтому рассмотрим более подробно базовый модуль 16V5-75 (см. рис. 1). Базовый модуль 16V5-75 обладает характеристиками: производительность (64 разряда) – 75 Гфлопс; производительность (32 разряда) – 140 Гфлопс; потребляемая мощность – 200 ВА; объем оперативной распределенной памяти – 1,25 Гбайт; количество ПЛИС решающего поля Virtex 5 XC5VLX110 – 16 шт.; количество эквивалентных вентилях в ПЛИС $8 \cdot 10^6$ шт.; тактовая частота – 250 МГц.; количество внешних LVDS каналов – 224 шт.; скорость межмодульного обмена – 22000 Мбит/сек.

Решающее поле базового модуля 16V5-75 выполнено на 16-ти ПЛИС Virtex 5 XC5VLX110-2FF1153 фирмы Xilinx, каждая из которых содержит около $8 \cdot 10^6$ эквивалентных вентилях. Микросхемы расположены в узлах двумерной решетки 4×4 и соединены между собой ортогональной системой связей по близкодействию посредством каналов LVDS. Такая система связей позволяет существенно упростить печатную плату и улучшить ее частотные характеристики. Данные между несмежными микросхемами передаются по транзитным каналам через промежуточные ПЛИС, используя систему ортогональных связей.

Базовый модуль имеет подсистему обмена информацией по быстрым LVDS-каналам с другими базовыми модулями, которая предназначена для передачи промежуточных результатов вычислений непосредственно из микросхем данного модуля в микросхемы решающего поля других базовых модулей. Единый тип интерфейса между ПЛИС как в пределах одного базового модуля, так и между различными базовыми модулями, обеспечивает одинаковые скорости обмена и снижает эффект границ в пределах общего схемотехнического ресурса PBC.

К периферийным ПЛИС решающего поля подключены микросхемы динамических ОЗУ SDRAM типа DDR2, которые образуют распределенную память базового модуля 16V5-75. Общий объем распределенной памяти составляет 1,25 Гбайт. Распределенная память предназначена для хранения входных данных, промежуточных и конечных результатов обработки информации, а также фрагментов параллельных программ. Кроме этого на базовом модуле установлено еще 128 Мбайт памяти, подключенной к контроллеру базового модуля – КБМ.

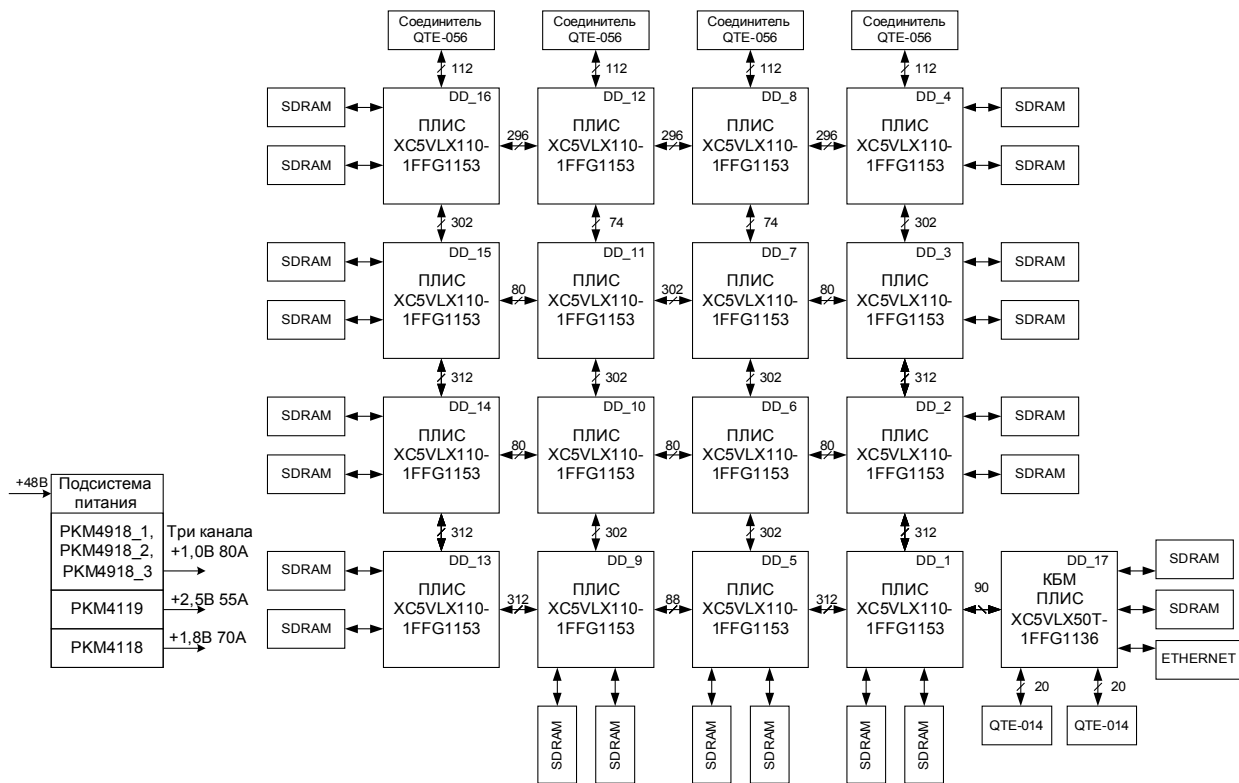


Рис. 1. Структура базового модуля 16V5-75

Основным назначением KBM являются функции управления подсистемами базового модуля, а также передача информации между базовым модулем и управляющим контроллером (ЭВМ типа IBM PC). Связь KBM с управляющим контроллером осуществляется посредством LVDS-каналов, а также с помощью канала Ethernet.

Подсистема электропитания расположена непосредственно на базовом модуле и содержит портативные преобразователи напряжения серии РКМ фирмы Ericsson, на которые подается напряжение +48В от источника первичного питания. 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.

Таким образом, базовый модуль 16V5-75 представляет собой мощный вычислительный узел производительностью свыше 75 Гфлопс. На основе БМ 16V5-75 могут строиться вычислительные блоки, содержащие от одного до восьми базовых модулей производительностью от 75 до 600 Гфлопс. В то же время базовый модуль обладает достаточной автономностью и может легко комплексоваться с персональным компьютером типа IBM PC в качестве ускорителя и использоваться при решении различных задач.

РАБОЧАЯ СТАНЦИЯ PBC-0.2-PC И ВЫЧИСЛИТЕЛЬНЫЙ БЛОК PBC-0.2-ВБ

На основе базового модуля 16V5-75 в рамках проекта разработаны рабочая станция PBC-0.2-PC и вычислительный блок PBC-0.2-ВБ производительностью 300 Гфлопс. Основу этих изделий составляет многопроцессорный вычислительный блок, который включает четыре базовых модуля, соединенных между собой в единый вычислительный ресурс быстрыми каналами LVDS и посредством коммутатора Ethernet.

Рабочая станция PBC-0.2-PC является представителем семейства PBC и предназначена для решения прикладных задач проектирования изделий микроэлектроники, управления в реальном времени сложными объектами, моделирования сложных технических и природных объектов и процессов, построения систем мониторинга, дистанционного зондирования, томографии и др.

Вычислительный блок PBC-0.2-ВБ практически полностью повторяет архитектуру рабочей станции PBC-0.2-PC, однако конструкции этих изделий значительно отличаются. Конструктивные отличия определяются назначением этих изделий: рабочая станция – это настольный вариант вычислительной системы, предназначенный для автономного использования, а вычислительный блок – это встраиваемый вариант, предназначенный для комплектования стоек СТ-1P в составе представителей семейства PBC-1P и PBC-5 и для создания суперЭВМ различных конфигураций. Вычислительный блок PBC-0.2-ВБ обладает теми же техническими параметрами, что и рабочая станция PBC-0.2-PC.

Реальная производительность PBC-0.2-PC и PBC-0.2-ВБ на задачах линейной алгебры, математической физики, цифровой обработки сигналов, символьной обработки составляет не менее 60% от пиковой производительности. Аппаратно-программные средства PBC-0.2-PC и PBC-0.2-ВБ позволяют динамически перестраивать архитектуру в процессе решения задачи на двух уровнях: программном (на уровне команд элементарных процессоров и каналов распределенной памяти), обеспечивающем высокую скорость реконфигурации системы на задачи из данного класса; схемотехническом (на уровне логических ячеек ПЛИС), обеспечивающем модернизацию системы команд элементарных процессоров и высокую удельную производительность системы при переходе на задачи различных классов.

Архитектурные отличия PBC-0.2-ВБ от PBC-0.2-PC заключаются в особенностях соединения базовых модулей. В PBC-0.2-PC базовые модули соединяются в кольцо, а в PBC-0.2-ВБ базовые модули имеют выходы за пределы корпуса блока с целью комплексирования нескольких вычислительных блоков PBC-0.2-ВБ в единую вычислительную структуру в составе стойки СТ-1Р. Во всем остальном архитектура PBC-0.2-ВБ совпадает с архитектурой PBC-0.2-PC.

Управляющий контроллер, входящий в состав обоих изделий, представляет собой персональный компьютер типа IBM PC, встроенный в корпус PBC-0.2-PC или PBC-0.2-ВБ и функционирующий под управлением операционной системы Windows XP. Управляющий контроллер предназначен для: управления всеми компонентами PBC-0.2-PC или PBC-0.2-ВБ; хранения информации (исходной, промежуточной и результатов); трансляции параллельных программ с языков высокого уровня в исполняемые коды; тестирования и диагностики; подключения периферийного оборудования; подключения к локальным и глобальным информационным сетям посредством Ethernet.

Системы охлаждения PBC-0.2-PC и PBC-0.2-ВБ имеют некоторые отличия, связанные с особенностями их назначения и эксплуатации.

РЕКОНФИГУРИРУЕМАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА PBC-1P

Реконфигурируемая вычислительная система PBC-1P предназначена для: оснащения научных центров с целью проведения исследований в области физики, химии, биологии, космоса, построения информационно-управляющих систем для управления потенциально опасными производствами, решения задач аэрокосмической, автомобильной промышленности и энергетики. Пиковая производительность PBC-1P составляет 1200 Гфлопс.

Основу PBC-1P составляет вычислительная стойка СТ-1Р, в которую монтируются четыре вычислительных блока PBC-0.2-ВБ, ПЭВМ, коммутатор Ethernet, система питания и система охлаждения. Четыре вычислительных блока PBC-0.2-ВБ объединяются с помощью быстрых LVDS-каналов в единый вычислительный ресурс, содержащий до 16-ти базовых модулей 16V5-75 с общей пиковой производительностью 1,2 Тфлопс. Межблочные связи являются продолжением межмодульных и, в свою очередь, продолжением связей между ПЛИС решающих полей базовых модулей. В целом подобная организация логических связей реализует в составе стойки СТ-1Р глобальный LVDS-канал передачи данных с единым темпом продвижения информации.

Для осуществления функций управления и мониторинга управляющие контроллеры вычислительных блоков соединяются между собой посредством коммутаторов EC0 и EC1 под общим управлением ПЭВМ, посредством которых можно установить прямые связи с любым из базовых модулей в составе стойки, минуя управляющие контроллеры блоков.

Встраиваемая ПЭВМ представляет собой промышленный компьютер типа IBM PC, встраивается в стойку СТ-1Р и предназначена для управления работой ресурсов вычислительной системы PBC-1P, а также для подключения СТ-1Р к локальным и глобальным информационным сетям посредством Ethernet.

РЕКОНФИГУРИРУЕМАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА PBC-5

Старшим представителем семейства PBC является изделие PBC-5 с пиковой производительностью 6000 Гфлопс, которое предназначено для использования в научно-исследовательских центрах с целью проведения исследований в области молекулярного моделирования, создания принципиально новых лекарственных препаратов и материалов нового поколения, построения информационно-моделирующих систем для долгосрочного прогнозирования погоды, моделирования задач сейсмоактивности, томографических исследований приповерхностных слоев Земли акустическими и электромагнитными волнами.

PBC-5 содержит пять стоек СТ-1Р, управляющую ЭВМ (УЭВМ), систему питания, систему охлаждения. К управляющей ЭВМ подключаются монитор, клавиатура, ручной манипулятор («мышь») и другие периферийные устройства. Основным вычислительным ресурсом PBC-5 являются пять стоек СТ-1Р, которые соединяются между собой посредством коммутаторов Ethernet с использованием сетевых технологий под общим управлением УЭВМ. Такое построение аппаратных средств позволяет максимально унифицировать оборудование, поскольку вычислительная часть PBC-5 строится из типовых узлов, уже использовавшихся для построения других представителей семейства. Развитая система информационных связей позволяет использовать различные варианты связей между вы-

числительными ресурсами PBC-5.

Управляющая ЭВМ представляет собой персональный компьютер типа IBM PC с типовым набором необходимых устройств и периферийного оборудования и используется для управления работой вычислительных ресурсов пяти стоек СТ-1Р, управления системами питания и охлаждения, подключения периферийных устройств, а также для подключения PBC-5 к локальным и глобальным информационным сетям.

РЕКОНФИГУРИРУЕМЫЕ УСКОРИТЕЛИ ПЕРСОНАЛЬНОГО КОМПЬЮТЕРА РУПК-50 И РУПК-25

Основным назначением реконфигурируемых ускорителей РУПК-25 и РУПК-50 является наращивание возможностей персональных компьютеров типа IBM PC при решении вычислительно трудоемких задач.

РУПК-50 и РУПК-25 предоставляют пользователю дополнительный вычислительный ресурс с пиковой производительностью свыше 50 и 25 Гфлопс соответственно. Реальная производительность РУПК-50 и РУПК-25 на задачах линейной алгебры, математической физики, цифровой обработки сигналов, символьной обработки превышает 60% от пиковой производительности. Аппаратно-программные средства позволяют динамически перестраивать архитектуру РУПК-25 и РУПК-50 в процессе решения задачи на программном и схемотехническом уровнях.

РУПК-50 и РУПК-25 имеют одинаковую структуру, содержат одинаковые узлы и различаются только типом базового модуля. РУПК-50 (РУПК-25) содержит: базовый модуль 16V5-50 (16S3-25) с устройством сопряжения с персональным компьютером; блок питания AC-DC; панель индикации; систему охлаждения.

Решающее поле базового модуля 16V5-50 выполнено на 16-ти ПЛИС Virtex 5 XC5VLX110-1FFG1153, а базовый модуль 16S3-25 – на 16-ти ПЛИС Spartan 3 XC3S4000-5FG1156 фирмы Xilinx. Распределенная память реализована на 24-х микросхемах SDRAM типа DDR2 и имеет общий объем 1,5 Гбайта. В отличие от базового модуля 16V5-75 базовые модули 16V5-50 и 16S3-25 не содержат каналов связи с другими базовыми модулями. Производительность базовых модулей 16V5-50 и 16S3-25 составляет 50 и 25 Гфлопс соответственно. Обмен информацией РУПК-50 и РУПК-25 с персональным компьютером по данным и управлению осуществляется с помощью ПЛИС КБМ, используя LVDS- или Ethernet-каналы.

СИСТЕМНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Программирование PBC отличается от программирования MBC традиционной архитектуры и его можно условно разделить на две составляющие: программирование структурное, которое создает необходимые вычислительные структуры в поле логических ячеек ПЛИС, и программирование процедурное – программирование в традиционном смысле, заключающееся в организации вычислительного процесса в PBC. При этом программирование вычислительных структур вызывает у пользователей наибольшие трудности [3, 4, 7, 8]. Это связано с тем, что пользователи традиционно привыкли программировать только организацию вычислительного процесса, опираясь на неизменяемую аппаратную поддержку средств вычислительной техники, в то время как для программирования вычислительных структур PBC требуется совершенно другая квалификация, а именно - квалификация схемотехника.

Целью создания программного комплекса средств разработки прикладных программ семейства PBC является предоставление пользователю возможностей, которые позволяют создавать программы без привлечения специальных знаний в области схемотехники ПЛИС и по сложности будут приближены к обычному программированию для микропроцессоров и многопроцессорных ЭВМ.

Создаваемое в рамках проекта системное и прикладное программное обеспечение включает в себя: программный комплекс средств разработки прикладных программ, средства администрирования вычислительных ресурсов семейства PBC и служебные программы и драйверы.

Программный комплекс средств разработки прикладных программ семейства PBC содержит: язык ассемблера Argus v.3.0; язык структурно-процедурного программирования высокого уровня COLAMO v.2.0; интегрированную среду разработки аппаратно-программных решений прикладных задач Argus IDE v.3.0, единую для всех представителей семейства PBC и поддерживающую языки программирования Argus v.3.0 и COLAMO v.2.0; отладчик параллельных программ на базовых модулях семейства PBC, поддерживающий межмодульные связи; программный интерфейс доступа к вычислительным ресурсам всех представителей PBC из различных сред программирования; среду разработки вычислительных структур для синтеза масштабируемых параллельно-конвейерных процедур, оперирующую библиотекой схемных решений (IP-ядер).

Создаваемое в рамках проекта системное и прикладное программное обеспечение должно обеспечить потенциальным пользователям удобство программирования сложных практических задач на семействе PBC.

Одновременно с разработкой семейства PBC создается и будет развиваться инфраструктура для обучения потенциальных пользователей и организации доступа к библиотекам масштабируемых IP-ядер для различных предметных областей с целью их широкого продвижения на рынке высокопроизводительной вычислительной техники. За счет использования возможностей сервисов глобальной компьютерной сети Internet планируется обеспечить широкий доступ потенциальных пользователей к информационным ресурсам и программным продуктам, созданным в результате выполнения проекта.

ЗАКЛЮЧЕНИЕ

Описанное в статье семейство РВС является принципиально новым направлением развития высокопроизводительной вычислительной техники. В качестве основного вычислительного элемента в РВС используются не универсальные микропроцессоры, а ПЛИС. Это дает возможность пользователям создавать в базовой архитектуре РВС виртуальные специализированные вычислители, структура которых адекватна структуре решаемой задачи, что, в свою очередь, обеспечивает высокую эффективность вычислений и близкий к линейному рост производительности при наращивании вычислительного ресурса. Семейство РВС с динамически перестраиваемой архитектурой на основе ПЛИС предназначено для решения вычислительно трудоемких задач, в том числе и большого размера, решение которых на МВС традиционной архитектуры либо требует недопустимо больших временных затрат, либо вообще невозможно.

ЛИТЕРАТУРА:

1. Аладышев О.С., Дикарев Н.И., Овсянников А.П. и др. СуперЭВМ: области применения и требования к производительности // Известия ВУЗов. Электроника, 2004. - №1. - С.13-17.
2. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. - С.-Петербург: «БХВ-Петербург», 2002. - 599 с.
3. Каляев А.В., Левин И.И. Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. - М.: Янус-К, 2003. - 380 с.
4. Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И. Реконфигурируемые мультимедийные вычислительные структуры / Под общ. ред. И.А.Каляева. - Ростов/Д: Издательство ЮНЦ РАН, 2008. - 320 с.
5. Каляев И.А., Левин И.И. Многопроцессорные вычислительные системы (суперкомпьютеры): состояние и перспективы // Вестник компьютерных и информационных технологий. - М.: Машиностроение, 2004. - №5. - С. 25-44.
6. Беседин И.В., Дмитренко Н.Н., Каляев И.А., Левин И.И., Семерников Е.А. Семейство базовых модулей для построения реконфигурируемых многопроцессорных вычислительных систем со структурно-процедурной организацией вычислений // Материалы Всероссийской научной конференции «Научный сервис в сети Интернет: технологии распределенных вычислений», г. Новороссийск. - М.: Издательство Московского университета, 2006. - С. 47-49.
7. Дордопуло А.И., Каляев И.А., Левин И.И., Семерников Е.А. Семейство многопроцессорных вычислительных систем с динамически перестраиваемой архитектурой // Материалы Четвертой Международной научной молодежной школы «Высокопроизводительные вычислительные системы». - Таганрог: Изд-во ТТИ ЮФУ, 2007. - С. 68-74.
8. Левин И.И. Язык параллельного программирования высокого уровня для структурно-процедурной организации вычислений // Труды Всероссийской научной конференции. - М.: Изд-во МГУ, 2000. - С.108-112.