

# СУПЕРЭВМ

## НА ОСНОВЕ АРХИТЕКТУРНОЙ ПЛАТФОРМЫ "ЭЛЬБРУС"

**Создание высокопроизводительных вычислительных систем – необходимое условие развития любого государства, выбирающего высокотехнологический путь развития. Это вопрос даже не столько экономический (что дешевле – купить или сделать самим), сколь затрагивающий проблему национальной безопасности и глобальной конкурентоспособности. Поэтому любые проекты, направленные на создание отечественной высокопроизводительной вычислительной техники, требуют особого, самого пристального внимания. Один из возможных подходов к решению проблемы предлагают авторы данной статьи.**

Россия, как страна, претендующая на роль одного из мировых лидеров, должна проектировать и изготавливать собственные компьютеры сверхвысокой производительности (суперкомпьютеры или суперЭВМ). Во все времена суперЭВМ во многом определяли и определяют стратегические приоритеты государства. В середине двадцатого столетия в СССР это хорошо понимали, когда закладывали институты по проектированию вычислительной техники, например ИТМ и ВТ или ИНЭУМ. Задача разработки вычислительной техники высокой производительности ставилась перед коллективом ученых и инженеров ИТМ и ВТ под руководством академика С.А.Лебедева [1]. Аналогичные задачи решались другими творческими коллективами. В частности, выдающийся вклад в создание высокоскоростных арифметических устройств вычислительной техники внес коллектив под руководством М.А.Карцева [2].

В СССР высокопроизводительная вычислительная техника, в первую очередь, требовалась для обеспечения обороноспособности страны. Высшим достижением в советский период был разработанный для систем ПРО многопроцессорный вычислительный комплекс "Эльбрус-2". Уникальные для того времени характеристики комплекса, в первую очередь высокая производительность, были достигнуты благодаря ряду инновационных технологий в области организации вычислительного процесса. Ими стали многопроцессорная структура на общей памяти с аппаратной поддержкой ко-

А.Ким, к.т.н.; kim@mcst.ru  
В.Фельдман, к.т.н.; feld@mcst.ru

герентности данных во всех уровнях иерархии памяти, аппаратная поддержка языков высокого уровня, суперскалярное исполнение программ, оптимизирующий компилятор, эффективные механизмы операционной системы, развитый аппаратный контроль и многое другое.

Идея достижения сверхвысокой производительности не за счет повышения тактовой частоты процессора, а за счет архитектурной производительности нашла отражение при разработке микропроцессора "Эльбрус" и вычислительного комплекса (ВК) "Эльбрус-3М1". Эти работы были завершены в конце 2007 года в ЗАО "МЦСТ" коллективом разработчиков, участвовавшим когда-то в разработке комплекса "Эльбрус-2", а также создавшим вычислительный комплекс "Эльбрус-90микро" [3, 4].

Предложенные решения были апробированы уже в предыдущих проектах, некоторые реализованы и запатентованы в процессе разработки. Вот некоторые из них:

- интеграция устройств, обрабатывающих целочисленные и вещественные числа на общем поле регистров;
- спекулятивный счет альтернатив в условных ветвлениях;
- реализация циклов путем проекции их элементов на широкую команду (без использования векторных операций);
- специальные средства предварительной подкачки данных.

Завершающий вывод государственной комиссии по приемке опытно-конструкторской работы: "Микропроцессор "Эльбрус" и ВК "Эльбрус-3М1" на его основе, их архитектура, общее программное обеспечение являются оригинальной отечественной разработкой, отражающей современные тенденции развития вычислительной техники. По архитектурно-логическим и программным решениям ВК "Эльбрус-3М1" находится на современном мировом уровне, а по ряду решений превосходит его".

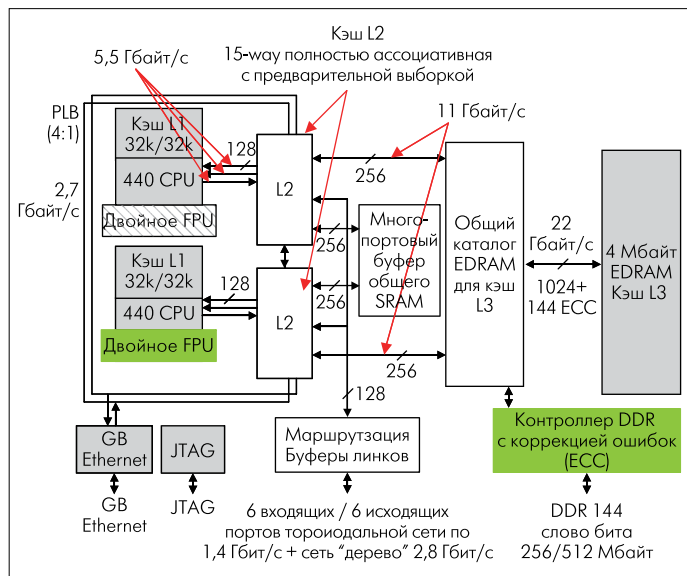
Полученные характеристики позволяют успешно использовать микропроцессор "Эльбрус" и его последующие модификации для построения отечественных суперЭВМ. За десять лет планируется увеличить производительность микропроцессоров более чем на два порядка (табл.1), в первую очередь – за счет многоядерности и увеличения тактовой частоты при переходе на передовые технологические нормы.

**Таблица 1. Планы в ОАО ИНЭУМ им. И.С.Брука и ЗАО "МЦСТ" по развитию линии микропроцессоров "Эльбрус"**

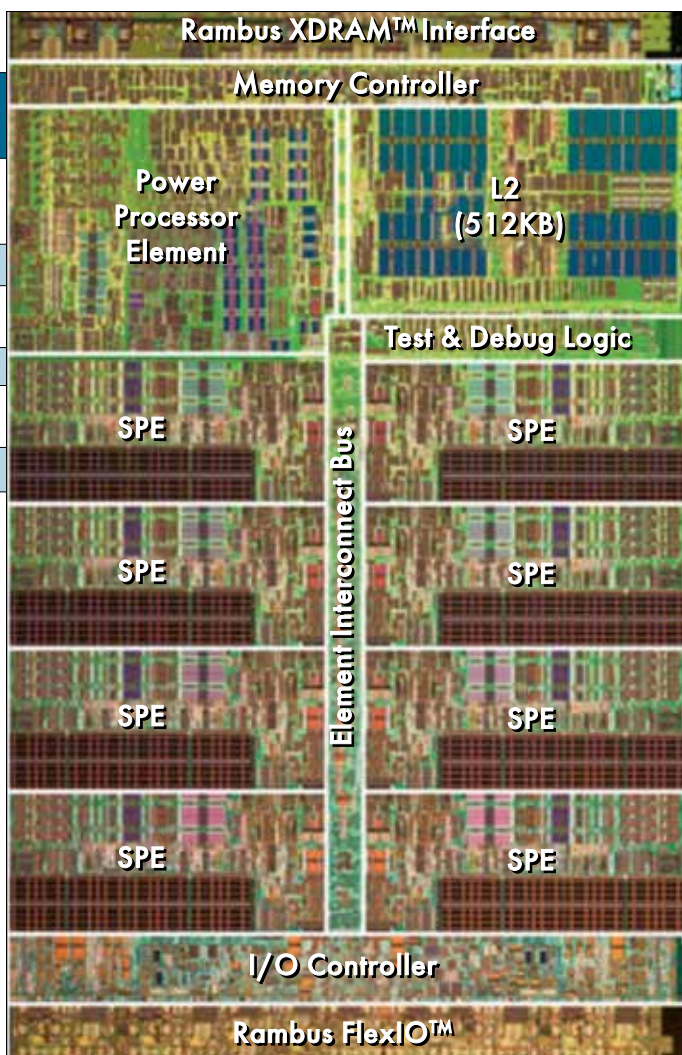
| Микро-процессор                | "Эльбрус" | "Эльбрус-2С" | "Эльбрус-4С" | "Эльбрус-8С" | "Эльбрус-16С" |
|--------------------------------|-----------|--------------|--------------|--------------|---------------|
| Технологическое разрешение, нм | 130       | 90           | 65           | 45           | 32            |
| Год выпуска                    | 2007      | 2010         | 2012         | 2015         | 2018          |
| Число процессорных ядер        | 1         | 2            | 4            | 8            | 16            |
| Частота, МГц                   | 300       | 600          | 1000         | 2000         | 3000          |
| Производительность, GFlops     | 4,8       | 19,2         | 64           | 256          | 768           |
| Мощность, Вт                   | 6         | 16           | 25           | 60           | 95            |

### ЭЛЕМЕНТАРНАЯ БАЗА СОВРЕМЕННЫХ СУПЕРЭВМ

Чтобы доказать, что микропроцессоры серии "Эльбрус" можно успешно использовать при построении суперЭВМ, рассмотрим элементную базу современных суперЭВМ, входящих в первую десятку списка TOP500 самых высокопроизводительных суперЭВМ мира, выпускаемого с 1993 года и обновляемого дважды в год [5]. Более 85% суперкомпьютеров, представленных в списке TOP500, базируются на микропроцессорах с архитектурой Intel IA-32 (x86). Однако первые места принадлежат суперкомпьютерам фирмы IBM: Roadrunner [6] и Blue Gene [7], в которых наряду с микропроцессорами архитектуры x86 (Opteron фирмы AMD в Roadrunner) применяются микропроцессоры IBM (двухъядерные PowerPC и четырехъядерные PowerPC450 в системах Blue Gene, PowerXCell 8i – в Roadrunner). Но если в Roadrunner используются обычные коммерческие микропроцессоры, для суперкомпьютера Blue Gene специально разработаны заказные микросхемы (ASIC). Так, в СБИС процессора PowerPC440 на одном



**Рис. 1. Блок-схема системы на кристалле для суперкомпьютера Blue Gene/L**



**Рис.2. Топология микропроцессора Cell**

кристалле размещены два процессорных ядра PowerPC 440, кэш-память третьего уровня, контроллеры оперативной памяти, GB Ethernet и высокоскоростных линков для связи с другими процессорами в системе (рис.1).

История построения суперЭВМ последних десяти лет показывает примеры успешного использования как коммерческих, так и заказных микропроцессоров. Нет явных доказательств правильности того или иного подхода с научной или технической точки зрения. Возможно, одним из критериев выбора является коммерческий – соотношение производительность/стоимость. Примеры проектов Blue Gene и Roadrunner демонстрируют, что при выборе элементной базы предпочтение отдается в первую очередь тем микропроцессорам, в которых реализована эффективная поддержка вычислений с плавающей запятой, так как именно эти вычисления составляют основу задач, которые решаются с помощью суперЭВМ. Большое число блоков обработки вещественных чисел, мощные каналы работы с памятью для поддержки этих вычислений и мощные каналы ввода-вывода для организации высокоскоростной коммуникационной сети – вот главные составляющие при выборе микропроцессора.

Микропроцессоры с архитектурой Cell Broadband Engine (рис.2), современная реализация которой, PowerXCell 8i, используется в суперЭВМ Roadrunner, обладают перечисленными выше качествами. Cell состоит из процессорного элемента на основе архитектуры Power (PPE), выступающего в качестве контроллера для восьми основных вычислительных блоков Synergistic Processing Elements (SPE). Для взаимодействия PPE и SPE, а также для обмена данными с контроллером памяти и I/O-интерфейсами служит внутренняя шина Element Interconnect Bus (EIB).

Процессоры PowerXCell 8i [8] предназначены специально для серверных систем. Основные их отличия от предшественников – технология уровня 65 нм и переход на более распространенной сегодня стандарт оперативной памяти DDR2 вместо Rambus XDR, которая поддерживалась оригинальными Cell. Объем доступной оперативной памяти возрос с 2 до 32 Гбайт. Изменения коснулись и блока вычислений с плавающей запятой, которые существенно увеличили свою производительность за счет полной конвейеризации. Теперь каждый SPE процессора PowerXCell 8i способен выполнять операцию с плавающей точкой с удвоенной точностью (т.е. с 64-разрядными операндами) за один такт. При тактовой частоте 3,2 ГГц процессор PowerXCell 8i обладает пиковой производительностью 102,4 GFlops (без учета возможностей ядра PPE).

Микропроцессор "Эльбрус-2С" (рис.3) по заложенным в него возможностям подобен микропроцессору Cell. Он включает 12 арифметических каналов ALU (по шесть на ядро), два контроллера памяти DDR2 (MC), когерентные каналы для межпроцессорной связи IPCC и канал ввода-вывода IOCC. Производительность микропроцессора "Эльбрус-2С" составит 19,2 GFlops – но на вычислениях с 32-разрядными операндами, т.е. с одинарной точностью. Это более чем на порядок меньше, чем у PowerXCell 8i, но при этом и тактовая частота "Эльбрус-2С" (600 МГц) ниже, чем у PowerXCell 8i, примерно в пять раз. Кроме того, в PowerXCell 8i в четыре раза больше процессорных элементов. Таким образом, с точки

зрения архитектурной производительности "Эльбрус-2С" не уступает микропроцессорам Cell.

Отметим, что соотношение производительность/мощность у процессора "Эльбрус-2С" составляет 1,2 GFlops/Вт. Это лишь в два раза хуже, чем у PowerXCell 8i – 2,2 GFlops/Вт (при 3,2 ГГц потребление порядка 92 Вт) для операций с одинарной точностью, учитывая только производительность вычислительных ядер SPE. Учитывая, что "Эльбрус-2С" использует технологию уровня 90 нм, а PowerXCell 8i – 65 нм, это очень неплохой показатель. А сегодня он не менее важен, чем производительность микропроцессора, используемого в качестве базового блока для построения суперЭВМ. Ведь суммарная мощность потребления современных суперЭВМ измеряется мегаваттами и, в конечном счете, ограничена. Отсюда вывод – чем больше соотношение производительность/мощность для одного элемента, тем больше будет суммарная производительность суперЭВМ при заданном лимите общей потребляемой мощности.

Очень важное качество архитектуры микропроцессоров серии "Эльбрус" – поддержка многопроцессорности. Каналы межпроцессорной связи IPCC позволяют непосредственно объединить до четырех микропроцессоров "Эльбрус-2С" в четырехпроцессорную систему, работающую с общей памятью при аппаратной поддержке когерентности памяти на всех уровнях иерархии. Специальный коммутатор когерентности позволит объединить до 16 микропроцессоров, работающих с общей памятью. Последнее очень важно, поскольку суперЭВМ, использующие общую память и аппаратную поддержку когерентности, на ряде задач демонстрируют большую эффективность, чем суперЭВМ, построенные по кластерной схеме с распределенной памятью и использующие механизмы MPI интерфейса [9]. Характерным примером является система Altix фирмы SGI. Разумеется, всегда можно подобрать задачи, для которых более эффективен MPI-интерфейс, особенно при большом числе процессоров, когда механизмы поддержки когерентности на аппаратном уровне становятся очень громоздкими и времязатратными.

Не менее важным достоинством многопроцессорной архитектуры с аппаратной поддержкой когерентности памяти является более простое программирование. Для программиста такая система выглядит как единое решающее поле с общей памятью, на котором распределением ресурсов ведаёт операционная система. В системах же с распределенной памятью программист должен явным образом распределять задачи по процессорам, реализовывать механизмы синхронизации при работе с общими данными. Все это на порядок усложняет программирование, требует сложной отладки, затягивает сроки и увеличивает стоимость проектов.

**ПРОБЛЕМЫ СОЗДАНИЯ И ИСПОЛЬЗОВАНИЯ СУПЕРЭВМ**

Создание и эксплуатация современных суперЭВМ сопряжены со множеством задач. Однако ключевой момент при создании суперкомпьютера – это элементная база и построение на ее

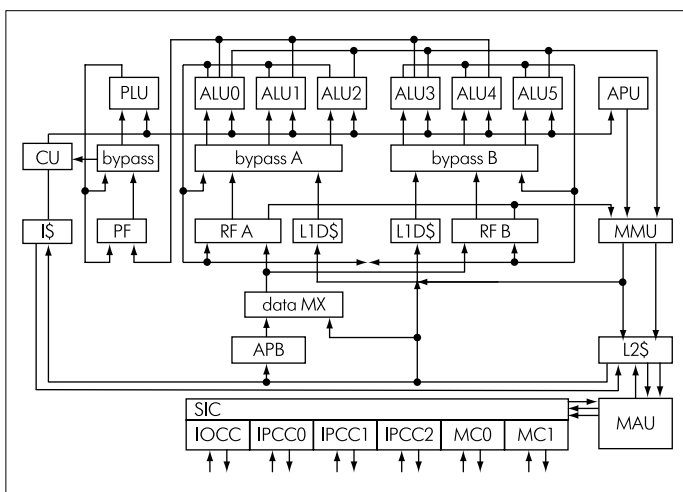


Рис.3. Структурная схема микропроцессора "Эльбрус-2С" (одно ядро)

основе вычислительных узлов. В дальнейшем эти узлы объединяются тем или иным способом, формируя собственно суперЭВМ. Именно так построены все современные высокопроизводительные вычислительные системы.

Для примера рассмотрим суперкомпьютер Roadrunner на момент его первого появления в списке TOP500 летом 2008 года\*. Эта система состояла из 278 стоек – шкафов IBM BladeCenter (каждый размером с холодильник). Она построена по гибридной схеме из 6120 двухъядерных процессоров AMD Opteron и 12240 процессоров IBM Cell 8i в специальных blade-модулях TriBlades, соединенных с помощью коммуникационной сети Infiniband. Установка занимает площадь приблизительно 560 м<sup>2</sup> и весит 250 т. Общее энергопотребление системы – 2,35 МВт, при этом энергоэффективность составляет 437 MFlops/Вт. Стоимость IBM Roadrunner составила 133 млн. долларов. Пиковая производительность суперкомпьютера – 1,376 PFlops, производительность на тесте Linpack – 1,026 PFlops.

Конструктивным элементом Roadrunner является так называемый гибридный вычислительный узел TriBlade. Он состоит из четырех blade-модулей – двух вычислительных модулей плат QS22, модуля LS21 и модуля расширения (Broadcom HT-2100). Каждый модуль QS22 содержит два процессора PowerXCell 8i и 8 Гбайт памяти, по 4 Гбайт на каждый процессор (возможно увеличение до 32 Гбайт). Пиковая производительность QS22 – 217 GFlops. LS21 содержит два двухъядерных процессора Opteron с 16 Гбайт памяти, по 4 Гбайт на ядро. Логически blade-модуль TriBlade (рис.4) состоит из четырех ядер Opteron, четырех процессоров PowerXCell 8i, 16 Гбайт памяти для Opteron и 16 Гбайт памяти для Cell. Процессоры Opteron выступают управляющими для PowerXCell 8i, каждому ядру Opteron соответствует один процессор Cell. Однако память модуля LS21 является общей для всего вычислительного узла – поддерживается архитектура с общей памятью с неоднородным и с аппаратной поддержкой когерентности (ccNUMA). Плата расширения соединяет QS22 через четыре канала PCI Express x8 с LS21. Также она обеспечивает подключение TriBlade к коммуникационной сети Infiniband 4x DDR.

Модули TriBlade монтируются в стойки трех функциональных типов. Стойка вычислительного узла содержит 12 модулей TriBlade. Стойка вычислительного узла с устройствами ввода/вывода также содержит 12 модулей TriBlade, но дополнительно включает коммуникационный сервер IBM System x3655, обеспечивающий доступ к файловой системе Panasas через каналы 10GB Ethernet. Стойка коммутации содержит 288-портовый Infiniband-маршрутизатор Voltaire ISR2012 и сервер System x3655.

\* Система Roadrunner постоянно модифицируется. К ноябрю 2008 года она включала 12960 процессоров PowerXCell 8i и достигала пиковой производительности 1,457 PFlops и производительности на тесте Linpack 1,105 PFlops. Энергоэффективность составляла уже 587 MFlops/Вт.

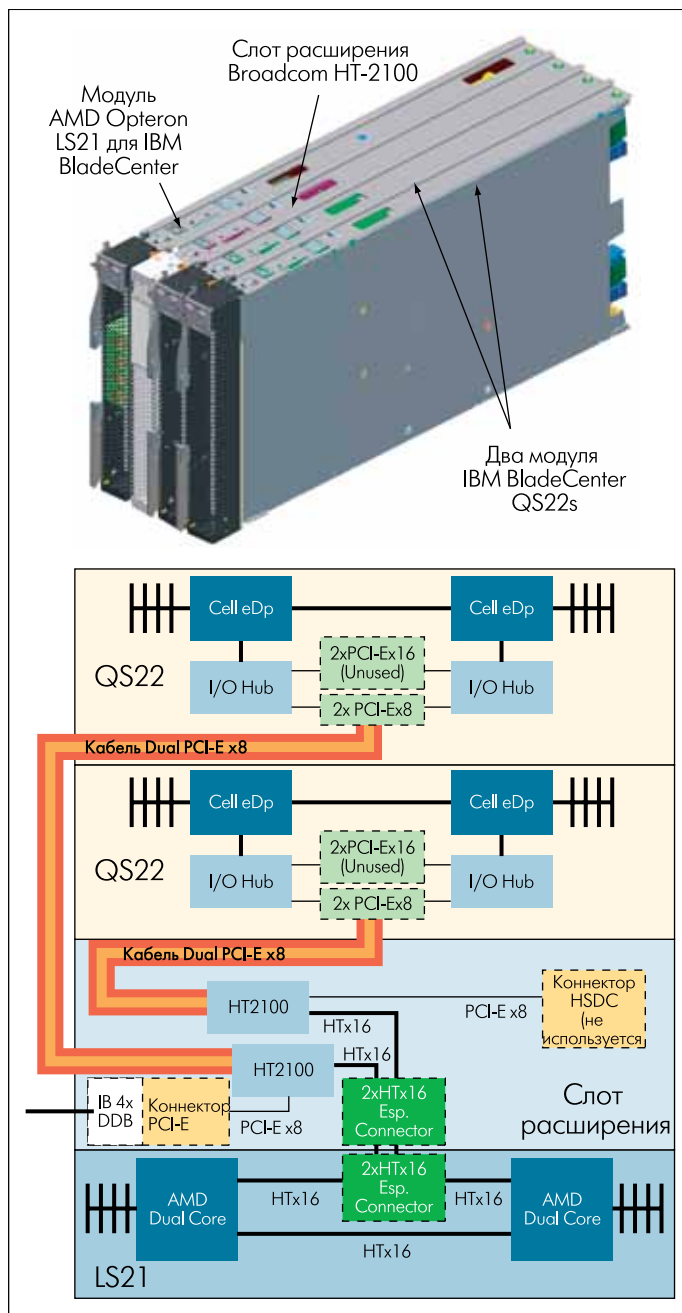
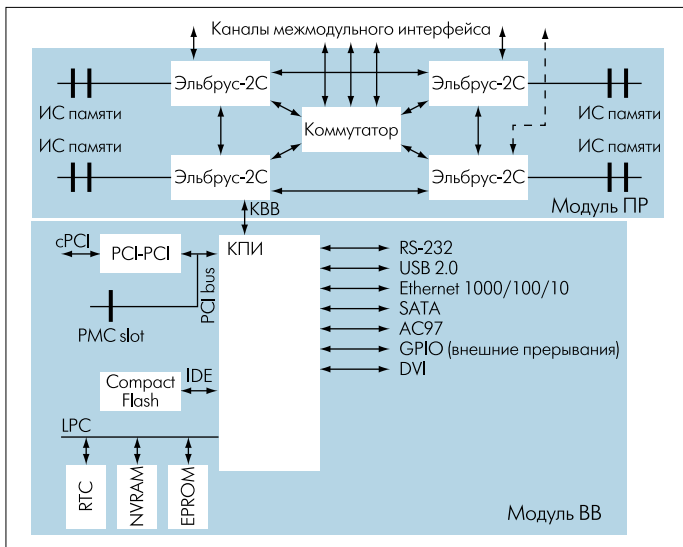


Рис.4. Структура blade-модуля TriBlade

Одна конструктивная единица Roadrunner – объединенный модуль (кластер), включающий три вычислительных стойки, 12 вычислительных стоек с серверами ввода/вывода и одну коммутационную стойку. Всего в объединенный модуль входит 180 модулей TriBlade, каждый из них подключен к отдельному порту коммутатора Infiniband (сеть первого уровня). Еще 8 портов служат для подключения серверов ввода/вывода, 96 портов – для связывания объединенных модулей в общую сеть (второго уровня).

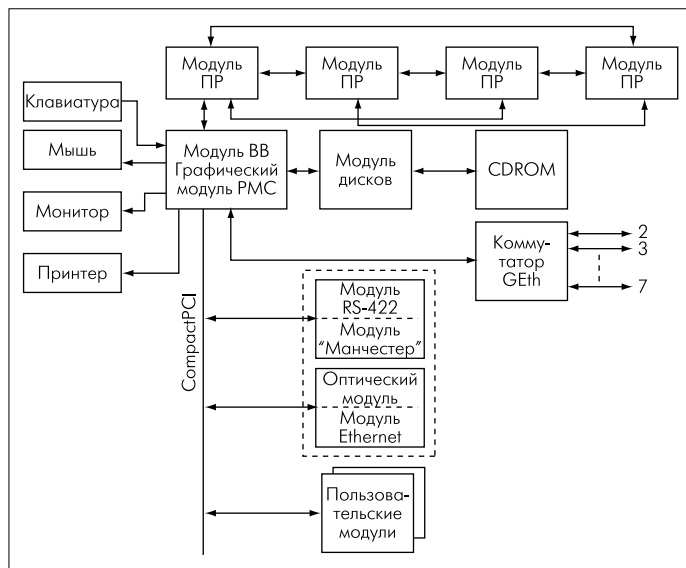
Кластер Roadrunner изначально был собран из 17 объединенных модулей, соединенных через восемь дополнительных (второго уровня) маршрутизаторов Voltaire ISR2012. Каждый объединенный модуль подсоединен через 12 каналов к каждому маршрутизатору, суммарно получается 96 соедине-



**Рис.5. Структурная схема процессорного модуля и модуля ввода-вывода сервера "Эльбрус-4"**

ний. Система Roadrunner работает под управлением операционной системы Fedora Linux и управляется при помощи программного обеспечения xCAT.

Серьезной проблемой является адаптация программного обеспечения (ПО) под новую платформу суперЭВМ. Уже создан огромный объем программного обеспечения для суперкомпьютеров. При росте производительности аппаратуры предъявляются серьезные требования к сохранению уже созданного программного обеспечения и к эффективности применения новых поколений суперкомпьютеров. Эти два требования противоречивы. Ведь уже существующее ПО, как правило, написано на языках высокого уровня (Fortran, C/C++ и т.п.) и во многих случаях не позволяет полностью использовать возможности новых поколений суперкомпьютеров. Только для того, чтобы продемонстрировать на задаче HPL (High Performance Linpack) производительность системы Roadrunner в 1,025 PFlops (при пиковой производительности



**Рис.6. Структурная схема 16-процессорного сервера "Эльбрус-4"**

1,325 PFlops, т.е. 77% от пиковой производительности), потребовалось дорабатывать этот совсем небольшой алгоритм (порядка 1 тыс. строк текста) в течение года. А современные программы для суперкомпьютеров состоят из сотен тысяч строк, и таких программ – тысячи. Не случайно в программе DARPA HPCS (High Productivity Computing Systems) уделяется специальное внимание сокращению трудоемкости создания и сопровождения программ как минимум в 10 раз для суперкомпьютеров петафлопного диапазона по сравнению с текущим уровнем. А это возможно только при условии, что программы не придется существенно перерабатывать при появлении новых поколений суперкомпьютеров

С другой стороны, использование уже существующих программ на новых поколениях суперкомпьютеров позволяет использовать их только на 10–15% от пиковой производительности, хотя эффективность использования может быть поднята почти на порядок.

**СУПЕРЭВМ НА БАЗЕ МИКРОПРОЦЕССОРОВ СЕРИИ "ЭЛЬБРУС"**

На примере суперкомпьютера Roadrunner мы показали, что столь высокопроизводительная вычислительная система может быть создана на базе стандартных элементов. Ее особенность – эффективные процессоры PowerXCell 8i и вычислительные узлы (серверы) на их основе. Сегодня, имея высокопроизводительные серверы, объединить их в большую многомашинную или многокластерную систему существенно проще, чем разрабатывать, например, сервер с тем же числом процессоров.

Межкластерные соединения в суперЭВМ выполняются обычно на стандартных интерфейсах – таких, например, как 10GB Ethernet, InfiniBand и др. Кластеры работают под управлением своей операционной системы и рассматриваются другими кластерами как внешние устройства. Как правило, задачи, решаемые на разных кластерах, не являются тесно связанными, однако между ними необходим обмен данными.

Таким образом, для создания суперЭВМ на отечественной элементной базе необходимы собственные высокопроизводительные процессоры и серверы на их основе. В качестве таковых можно использовать микропроцессоры серии "Эльбрус-С" и серверы ряда "Эльбрус". В частности, в ближайших планах компании "МЦСТ" – создание сервера "Эльбрус-4" на базе процессоров "Эльбрус-2С". Разработчики предлагают, как и в суперкомпьютере Roadrunner, использовать эти готовые высокопроизводительные серверы для создания суперЭВМ.

В серверах ряда "Эльбрус" используются подходы, которые давно применяются компанией "МЦСТ" для создания вычислительных систем:

- построение сервера как многопроцессорной системы, работающей на общую память с аппаратной поддержкой когерентности памяти;



- конструктивная реализация модулей и блоков по стандарту "Евромеханика";
- использование двух типов модулей: процессорных и ввода-вывода;
- применение средств автоматического распараллеливания вычислений на этапе компиляции программ;
- доработка механизмов операционной системы Linux для поддержки работы с архитектурой ccNUMA.

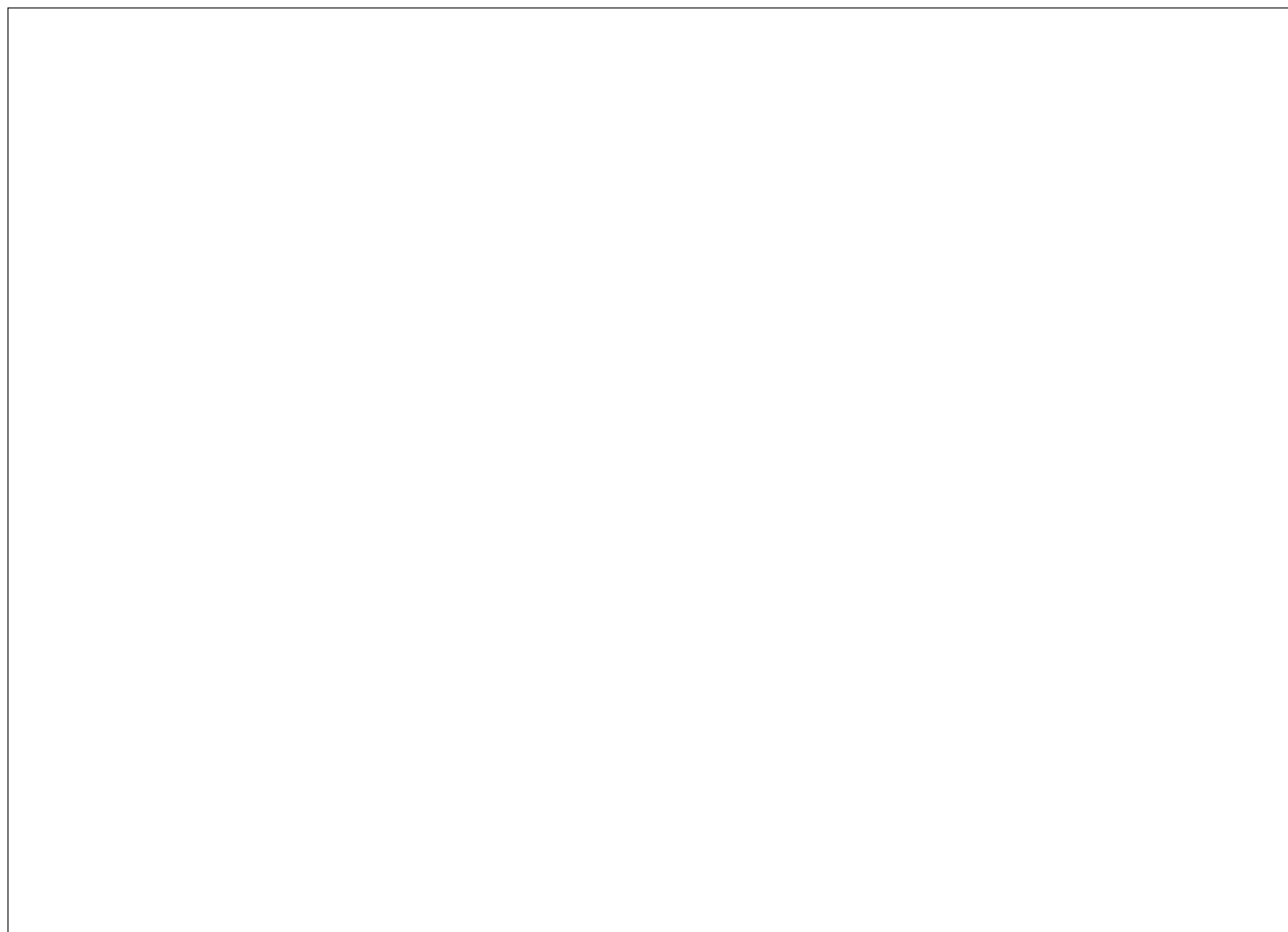
Архитектура ccNUMA подразумевает систему, где все процессоры работают с общей памятью, охвачены аппаратной поддержкой когерентности кэш-памяти, но с несимметричным доступом к различным доменам памяти. Вариант построения многопроцессорного комплекса с этой архитектурой выглядит более предпочтительным по сравнению с вариантом многомашинного построения сервера. В первую очередь, при реализации архитектуры ccNUMA достигаются более высокие характеристики суммарной производительности сервера. Кроме того, такое решение значительно упрощает программирование, а также позволяет сократить число модулей в сервере за счет снижения числа модулей ввода-вывода.

Сервер "Эльбрус-4" содержит 16 процессоров "Эльбрус-2С", размещенных на четырех процессорных модулях (рис.5). Четыре процессорных модуля, объединенные каналами межмодульного интерфейса, связаны с од-

**Таблица 2. Серверы "Эльбрус"**

| Вычислительные комплексы               | "Эльбрус 3М" | "Эльбрус-4" | "Эльбрус-5" | "Эльбрус-6" | "Эльбрус-7" |
|--|--------------|-------------|-------------|-------------|-------------|
| Год выпуска                            | 2007         | 2011        | 2013        | 2016        | 2019        |
| Производительность процессора (GFlops) | 4,8          | 19,2        | 64          | 256         | 768         |
| Число процессоров на сервере           | 2            | 16          | 64          | 64          | 64          |
| Производительность сервера (GFlops)    | 9,6          | 300         | 4000        | 16000       | 49000       |
| Число серверов/шкафов в супер-ЭВМ      | 1024/256     | 2048/256    | 2048/256    | 2048/256    | 2048/256    |
| Производительность суперЭВМ (PFlops)   | >0,01        | >0,6        | >8          | >32         | >100        |

ним модулем ввода-вывода (ВВ) через канал ввода-вывода КВВ. Из модуля ввода-вывода (рис.6) выходят несколько каналов, включая шину CompaqPCI, дисковый интерфейс, сеть Ethernet и стандартные интерфейсы для ввода и



отображения информации. Канал Ethernet соединен с восьмиканальным модулем коммутатора. На шине CompactPCI размещены модули-носители РМС-мезонинов, в которых могут быть установлены не требующиеся для суперЭВМ модули оптоэлектронных каналов, модуль интерфейса по ГОСТ 26765.52-87 ("Манчестер-II"), модули каналов RS-422 и Ethernet (сервер "Эльбрус-4" разрабатывается под определенные задачи, поэтому в нем имеются каналы и модули, которые для создания суперЭВМ не требуются).

Суммарная пиковая производительность сервера составит не менее 300 Gflops (при работе с 32-разрядными данными), суммарный объем оперативной памяти – не менее 32 Гбайт, пиковая суммарная пропускная способность оперативной памяти – 128 Гбайт/с, суммарная пропускная способность каналов межпроцессорного и межмодульного интерфейса – 60 Гбайт/с, пропускная способность канала ввода-вывода – 2 Гбайт/с, объем внешней памяти – не менее 300 Гбайт, потребляемая мощность – около 600 Вт.

Конструктивно модули, изображенные на рис.5, размещаются в корпусе "Евромеханика-6U" с системной шиной Compact PCI и 21 посадочным местом. Процессорные модули могут занимать от четырех до восьми мест. Это связано с решением проблемы снятия тепловыделения в этих модулях. По предварительным оценкам, максимальная выделяемая мощность составит порядка 100 Вт. В мире есть примеры решений снятия такой мощности с модуля формата 6U CompactPCI (COOL-H High Power 9-Slot Towers фирмы Hybricon [10]), но при этом, чтобы остаться в шаге установки модулей примерно 20 мм с воздушным охлаждением, используются специальные патентованные решения. Другой вариант решения проблемы тепла – это переход на больший шаг между процессорными модулями (полуторный или двойной) и установка на модулях мощных радиаторов.

Для примера, четыре процессорных модуля будут занимать шесть мест, одно место – модуль ВВ с установленным на нем графическим модулем, одно место – модуль коммутатора Ethernet и шесть мест – три модуля питания (суммарной мощностью 600 Вт). Остается еще семь посадочных мест, которые могли бы быть заняты дополнительными процессорными модулями. При использовании сервера "Эльбрус-4" в суперЭВМ, в вычислительную стойку таких же размеров, как в Roadrunner, можно будет установить восемь серверов "Эльбрус-4", получив 128 процессоров на одну стойку. У Roadrunner в одной стойке 72 процессора (из них 48 Cell-процессоров). В результате по удельной производительности на единицу объема "Эльбрус-4" уступает Roadrunner чуть более, чем в три раза. И это очень неплохой показатель, если учесть, что микропроцессор Cell имеет производительность на порядок выше производительности микропроцессора "Эльбрус-2С".

Важным вопросом является выбор каналов связи между процессорами и топология сети связи. Топология сети мо-

жет потребовать наличия хабов и/или коммутаторов. Поэтому, учитывая возможности ввода/вывода серверов "Эльбрус-4", для межкластерных связей можно использовать один из двух пригодных для этих целей интерфейсов – GB Ethernet или линки ввода/вывода, специально разработанные для этих серверов. Линк ввода/вывода – это полнодуплексный канал с пропускной способностью 1 Гбайт в каждом направлении, что значительно превышает возможности GB Ethernet. Межкластерную сеть можно построить и используя технологию InfiniBand. Для этого, как и в Roadrunner, необходимы специальные мосты для связи с сетью InfiniBand.

Общая максимальная мощность сервера "Эльбрус-4" составит около 600 Вт. Системная эффективность мощности сервера составит 0,5 GFlops/Вт (для 32-разрядной арифметики), что лишь в два раза хуже, чем в стартовой версии суперкомпьютера Roadrunner (0,437 GFlops/Вт при 64-разрядных данных). Такая суперЭВМ может быть построена в 2011 году, параллельно с выпуском сервера "Эльбрус-4". В перспективе, при построении серверов серии "Эльбрус" на последующих микропроцессорах ряда "Эльбрус" (см. табл.1), можно будет в ближайшее десятилетие достигнуть производительности суперЭВМ до 100 PFlops (табл.2).

#### ЛИТЕРАТУРА

1. **Лебедев С.А.** Электронная счетная машина. – Правда, 4 декабря 1955 г.
2. **Карцев М.А., Брик В.А.** Вычислительные системы и синхронная арифметика. – М.: Радио и Связь, 1981.
3. **Ким А.К.** Развитие архитектуры вычислительных комплексов серии "Эльбрус". – Материалы конференции "Перспективы развития высокопроизводительных архитектур. История, современность и будущее отечественного компьютеростроения". – Сборник научных трудов ИТМ и ВТ, 2008, Вып.1.
4. **Фельдман В.М.** Вычислительные комплексы "Эльбрус-90микро". – Информационные технологии, 2005, №1, с. 17–26.
5. <http://www.top500.org>
6. **A.Komornicki, G.Mullen-Schulz, D.Landon.** Roadrunner: Hardware and Software Overview. IBM Redpaper. – IBM, 2009. [www.redbooks.ibm.com](http://www.redbooks.ibm.com)
7. **Nicholas Allsopp et al.** Unfolding the IBM Eserver Blue Gene Solution. IBM Redbook. – IBM, 2006. [www.redbooks.ibm.com](http://www.redbooks.ibm.com)
8. **PowerXCell 8i CMOS SOI 65 nm Hardware Initialization Guide.** – IBM, December 2008.
9. **P.Lammers, G.Wellein et al.** Application Performance of Modern Number Cruncher. – CSAR Focus, 2004. edition 12. [www.csar.cfs.ac.uk/about/csarfocus/focus12/application\\_performance.pdf](http://www.csar.cfs.ac.uk/about/csarfocus/focus12/application_performance.pdf)
10. <http://www.hybricon.com>