РОССИЙСКАЯ АКАДЕМИЯ НАУК

РОССИЙСКИЙ ФОНД ФУНДАМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

«ОТКРЫТЫЕ СИСТЕМЫ»

Третий Московский суперкомпьютерный форум

ТЕЗИСЫ ДОКЛАДОВ

Москва, 2012 год

Задача национального масштаба

В современном мире преимущества получает тот, кто сумеет выловить крохи информации из океана данных: идет ли речь о выявлении сведений о готовящемся теракте, отфильтрованных в социальных сетях, или об обнаружении нерегулярных событий в телеметрии территориально размещенных сенсоров — везде требуются суперкомпьютеры. Идеей создания таких суперкомпьютеров озаботились уже во многих странах, однако для этого требуется решить множество серьезных проблем, причем не только технологических, — сегодня ни одна ИТ-компания и даже страна еще не может заявить, что обладает всем пакетом технологий для суперкомпьютера мощностью 1 EFLOPS. Для создания такой системы необходимо комплексное решение задач энергопотребления, повышения эффективности работы с распределенной памятью, распараллеливания приложений и т. п.

На Третьем Московском суперкомпьютерном форуме обсуждались актуальные тенденции в сфере создания аппаратно-программного обеспечения и эксплуатации высокопроизводительных распределенных компьютерных инфраструктур и была организована площадка для обмена результатами исследований, а главное — МСКФ стал очередным шагом к консолидации работ отечественных специалистов, опирающихся на национальные научно-технические разработки при выполнении проектов, соответствующих мировому уровню. В программу форума вошли доклады представителей практически всех организаций, занятых сегодня в суперкомпьютерной тематике, включая блок сообщений по опыту применения высокопроизводительных систем в национальной экономике, а также по организации подготовки кадров для ИТ-индустрии. В целом доклады МСКФ свидетельствуют о перспективности тематики форума и необходимости регулярного проведения аналогичных встреч представителей всех команд отечественных исследователей и разработчиков, принимающих участие в создании экзафлопсных технологий.
Создание суперкомпьютера экзафлопсного уровня — задача национального масштаба, решаемая в развитых странах на уровне федеральных программ, поэтому для координации усилий отечественных разработчиков и концентрации имеющихся в России ресурсов требуется, в частности, координирующее участие государства, участие промышленных предприятий, что позволит оптимизировать ресурсы на исследования и разработки в данной области.

— Юрий Гуляев, председатель оргкомитета МСКФ-2012

УДК 519.6, 519.7, 004.2, 004.3
ББК 32.973

Оргкомитет Третьего Московского суперкомпьютерного форума выражает признательность за поддержку Российскому фонду фундаментальных исследований (грант 12-07-06085-г), Отделению нанотехнологий и информационных технологий РАН, компаниям «АйТи», «ИТ-Экспо», «РСК Технологии», «Т-Сервисы» и IBM.

Тематика МСКФ-2012:
Пленарная сессия. Тенденции и перспективы суперкомпьютерной индустрии
Секция. Экзафлопсные технологии
Секция. Суперкомпьютерные архитектуры
Секция. Кадры для индустрии высокопроизводительных вычислений
Секция. Высокопроизводительные системы для решения практических задач

Тезисы докладов Третьего Московского суперкомпьютерного форума (Москва, 1 ноября 2012 г.). / [Под ред. Волкова Д.В.]. — М.: «Открытые системы», 2012. —48 с.

В сборник трудов включены доклады Третьего Московского суперкомпьютерного форума, прошедшего 1 ноября 2012 года в Москве в рамках XXIII российской выставки Softool-2012. Целями форума были обсуждение актуальных вопросов в сфере создания аппаратно-программного обеспечения и эксплуатации высокопроизводительных распределенных компьютерных инфраструктур, обмен результатами исследований и консолидация работ отечественных специалистов, опирающихся на национальные научно-технические разработки при выполнении проектов, соответствующих мировому уровню. Материалы сборника предназначены для научных сотрудников, преподавателей, аспирантов и студентов, интересующихся проблемами создания и эксплуатации суперкомпьютерных систем. Подробную информацию о МСКФ-2012 можно найти по адресу www.ospcon.ru.

Реализация в России любой крупной задачи, включая обеспечение национальной безопасности или функционирование платформы электронной демократии, невозможна сегодня без отечественных высокопроизводительных компьютерных систем. На Третьем Московском суперкомпьютерном форуме обсуждались вопросы создания аппаратно-программного обеспечения и эксплуатации суперкомпьютеров экзафлопсного уровня, проводился обмен результатами исследований с целью консолидации работ всех разработчиков. Доклады МСКФ свидетельствуют о необходимости регулярного проведения аналогичных форумов с участием представителей всех отечественных исследователей и разработчиков, создающих экзафлопсные суперкомпьютеры. Учитывая небывалую сложность и комплексность задач создания таких систем, целесообразно развертывание в отечественной университетской и академической среде групп информационного анализа. Следует также обратить внимание на подготовку кадров, способных проводить исследования фундаментального характера в области архитектуры и программного обеспечения компьютеров экзамасштабного уровня производительности.

Создание экзафлопсных технологий — задача национального масштаба, и для координации усилий всех разработчиков, концентрации имеющихся в России ресурсов и обеспечения региональной кооперации работ в этой области требуется участие федерального правительства, что позволит сократить время и затраты на исследования и разработки в суперкомпьютерной индустрии.

Организационный комитет

Гуляев Юрий Васильевич	академик РАН, член президиума РАН, директор ИРЭ им. В.А. Котельникова РАН, председатель оргкомитета
Волков Дмитрий Владимирович	гл. редактор журнала «Открытые системы. СУБД», зам. председателя оргкомитета
Бетелин Владимир Борисович	академик РАН, директор НИИСИ РАН
Воеводин Владимир Валентинович	чл.-корр. РАН, зам. директора НИВЦ МГУ
Иванников Виктор Петрович	академик РАН, директор Института системного программирования РАН
Корнеев Виктор Владимирович	д.т.н, зам. директора по науке ФГУП «НИИ "Квант"»
Кузьминский Михаил Борисович	к.хим.н., зам. зав. лаб. ИОХ РАН
Савин Геннадий Иванович	академик РАН, директор МСЦ РАН
Слуцкин Анатолий Ильич	к.т.н., научный руководитель по направлению суперкомпьютерных технологий ОАО «НИЦЭВТ»
Соловьев Вячеслав Петрович	д. физ.-мат. н., первый зам. директора ФГУП «РФЯЦ-ВНИИЭФ», директор ИТМФ
Сухомлин Владимир Александрович	д.т.н., профессор МГУ
Христов Павел Вячеславович	к.физ.-мат.н., вице-президент ЗАО «Открытые системы»
Фельдман Владимир Марткович	д.т.н., зам. ген. директора ОАО «ИНЭУМ им. И.С. Брука»
Черепенин Владимир Алексеевич	чл.-корр. РАН, зам. директора ИРЭ им. В.А. Котельникова РАН

ПЛЕНАРНАЯ СЕССИЯ.
ТЕНДЕНЦИИ И ПЕРСПЕКТИВЫ СУПЕРКОМПЬЮТЕРНОЙ ИНДУСТРИИ

На пути к экзафлопсному суперкомпьютеру: результаты, направления, тенденции

Эйсымонт Л.К. (verger-lk@yandex.ru), Горбунов В.С., Елизаров Г.С. — ФГУП «НИИ "Квант"», (Москва)

В области суперкомпьютерных технологий (СКТ) сегодня весьма актуальна экзафлопсная тематика. Она чаще ассоциируется с созданием суперкомпьютеров экзафлопсной производительности (10^{18} операций над вещественными числами в секунду, или 1 EFLOPS, или 1000 PFLOPS) и обеспечением энергетической эффективности вычислений в 50 Гфлопс/Вт, что в десятки раз выше современного уровня. При этом следует учесть, что экзафлопсная производительность — это не абстрактная пиковая характеристика аппаратуры, а характеристика, полученная при выполнении реальных задач пользователей. Именно такая цель стоит перед мировой суперкомпьютерной индустрией в текущем десятилетии, а до этого целью было достижение реальной производительности в 1 PFLOPS. Этот барьер был преодолен в 2008 году на суперкомпьютере Cray Jaguar (США). Работы этого направления продолжаются до сих пор, поскольку не все важные приложения допускают такой эффективный счет, и одно из важнейших направлений исследований — разработка новых численных алгоритмов.

Достижение высокой реальной производительности гораздо сложнее, чем пиковой, поскольку операции надо обеспечить операндами, а для значительной части задач, из-за характерной для них плохой пространственно-временной локализации, данные для операндов извлекаются не из быстрых кэш-памятей, а из оперативной памяти, длительность обращений к которой соизмерима с сотнями тактов процессора, что вызывает его простаивание. Другая проблема состоит в том, что операции реализуются в результате выполнения машинных команд, для которых на выборку из памяти и подготовку к выполнению заданных в них операций тоже тратятся время и энергия.

Проблема работы с памятью стала актуальной еще в прошлом десятилетии и получила название «стена памяти». Она решалась еще в петафлопсной программе DARPA HPCS (США) и аналогичных программах в Китае и Японии. В этих программах предполагалось применение ограниченных в распространении инновационных технологий, однако не отвергались и коммерчески доступные технологии, применение которых позволило быстро добиться первых успехов.

Обеспечение высокой реальной производительности на экзауровне еще сложнее, поскольку возникают новые требования и усиливаются ограничения современных элементно-конструкторских технологий. Кроме этого, создание экзафлопсных суперкомпьютеров рассматривается как часть более общей задачи создания «экзамасштабных технологий», используемых не только при создании гигантских компьютеров с экзафлопсной производительностью, но и петафлопсных суперкомпьютеров в одной стойке и терафлопсных серверов на одной плате (такая задача была сформулирована в США в новой программе DARPA UHPC, стартовавшей в 2010 году).

Программа DARPA UHPC существенно ориентирована на инновации, что явно указано в ее преамбуле — «…проекты на базе традиционных (коммерчески доступных) технологий не предлагать». Позже это требование было усилено в программе фундаментальных исследований DARPA OHPC, проводимой в интересах DARPA UHPC, работы в рамках которой отличаются новой организацией исполнителей — созданы четыре мощные группы из суперкомпьютерных и микроэлектронных компаний, национальных лабораторий и университетов. Сегодня наметилась тенденция вывода обсуждения ряда вопросов на международный уровень с целью вовлечения в работы специалистов из Европы и Азиатско-Тихоокеанского региона.

Проекты программы DARPA UHPC отражают свойственный Министерству обороны США (DoD) инновационный характер решения проблем, причем в сторону инновационности меняется также бывший традиционно эволюционным характер работ и в Министерстве энергетики США (DoE). В дополнение к инновационным работам ее лабораторий, в начале июля этого

года для поддержки проектов DARPA UHPC была запущена программа FastForward (компании Intel, NVIDIA, AMD и Whamcloud, разрабатывающая системы хранения данных). Эта программа рассматривается как первая фаза будущей крупной программы DoE 2013 года и запущена на фоне продолжающихся в DoE работ эволюционного направления развития суперкомпьютеров, включающего работы над «тяжелыми» процессорными ядрами (использование мощных и энергоемких коммерчески доступных универсальных многоядерных микропроцессоров, ускорителей и заказных коммуникационных сетей — линейка Cray XT/XE/XK) и над «легкими» процессорными ядрами, использующими большое количество не очень мощных и экономичных заказных микропроцессоров и сетей, специальных методов компоновки вычислительных узлов (линейка IBM BlueGene L/P/Q). Эти работы ведутся в Окриджской (ORNL) и Аргонской (ANL) лабораториях DoE.

Вместе с развертыванием работ по экзамасштабным технологиям в текущем десятилетии изменилась и общественная значимость работ в области суперкомпьютинга. Если раньше суперкомпьютеры рассматривались как стационарные установки для решения единичных вычислительных научно-технических задач, то сейчас усиливается потребность в их массовом применении для инженерных расчетов в промышленности. Одновременно приобрели значимость информационно-аналитические задачи в социально-экономической области и области управления. Таким образом, суперкомпьютерная тематика перестала быть экзотическим направлением, а стала жизненной необходимостью обеспечения развития экономики и промышленности развитых стран.

Основные цели эволюционных и инновационных программ прошлого десятилетия, в частности программы DARPA HPCS и близких к ней программ Китая и Японии, можно считать достигнутыми: созданы суперкомпьютеры IBM BG/Q (коммерческий вариант IBM Power 775), Cray Baker (коммерческий вариант Cray XE6/XK6), Cray Scorpio, Cray XMT2 (информационно-аналитические задачи), Tianhe-1A, K-компьютер. Напомним эти цели: эффективная реализация глобально адресуемой памяти для повышения в десятки раз продуктивности создания параллельных программ и повышения эффективности решения задач с хорошей пространственно-временной локализацией работы с памятью (до десяти раз, на тесте Linpack — не менее 2 PFLOPS) и задач с плохой пространственной локализацией (на 3–4 порядка, на тесте RandomAccess, на котором следовало достичь показателя 64 000 GUPS). Первый показатель был превзойден на японском K-компьютере (10 PFLOPS), но по второму был получен весьма «скромный» результат — 121 GUPS на K-компьютере и 117 GUPS на IBM BlueGene/P.

В середине июля 2012 года ситуация резко изменилась, специалисты IBM опубликовали результат в 1571,91 GUPS, что почти в 13 раз больше официального рекорда, и получен он был на суперкомпьютере IBM Power 775 (BG/Q), содержащем 1474 процессора. В этом эксперименте с тестом RandomAccess к каждому суперузлу была добавлена серверная плата Tesla C1060 с шестью GPU, что оказалось решающим для получения высокого результата. Такая гибридная конфигурация согласуется с концепцией создававшихся в рамках DARPA HPCS суперкомпьютеров, ориентированных на гибридную вычислительную модель программ, оборудование с раздельным доступом к данным и раздельно выполняемыми вычислениями, массовую мультитредовость и глобально адресуемую память.

Сегодня в ведущих суперкомпьютерных центрах США идет внедрение результатов «петафлопсных» программ прошлого десятилетия, в том числе и DARPA HPCS, как это и планировалось изначально.

В будущем увеличение суперскалярной многоядерности за счет улучшения микроэлектронных технологий уже не позволит 1000-кратно поднять производительность, поэтому базовая идея преодоления экзафлопсного барьера состоит в 1000-кратном увеличении параллелизма. Однако это означает возникновение серьезнейших проблем при разработке прикладного и системного ПО, обеспечении отказоустойчивости и самовосстанавливаемости, причем все это на фоне требования резкого снижения энергопотребления. Перспективные ключевые технологические решения: технологии 3D-компоновки, новые методы теплоотвода,

новая микроархитектура процессоров и DRAM-кристаллов памяти, внутрикристальные и межкристальные оптические коммуникационные сети.

Ключевая проблема прошлого десятилетия («стена памяти») сейчас перерастет в более крупную — «перемещения данных». Вынужденный рост параллелизма выполняемых операций и увеличение потока обращений к памяти для обеспечения параллелизма и толерантности к задержкам по памяти вступают в противоречие с допустимыми уровнями потребляемой энергии. Если не предпринимать специальных мер, то мощность потребления энергии экзафлопсной системой будет 150–200 МВт, что сопоставимо с мощностью атомной силовой установки современного многоцелевого авианосца. Это, одновременно, очень дорого в эксплуатации, плата за электроэнергию для такой системы будет составлять 100 млн долл. в год. По этой причине ставится задача не превысить уровня потребляемой энергии в 20 МВт.

На что уходит энергия? Около 70% расходов — это хранение данных в памяти и их перемещения, поэтому нужны решения по повышению локализации как данных при вычислениях, так и вычислений при данных, новые модели параллельных программ и новые технологии кристаллов памяти. Для достижения потребления в 20 МВт вместо 150–200 МВт надо не только решить проблему передачи и хранения данных, но и резко снизить накладные расходы на организацию параллельного выполнения огромного количества операций. Расходы на пересылку команд, дешифрацию и управление выполнением команд могут в десятки раз превышать затраты на выполнение собственно операций, кодируемых этими командами.

В области создания экзафлопсных технологий основные направления работ в текущем десятилетии будут задаваться программой DARPA UHPC создания экзамасштабных технологий:

- повышение энергоэффективности на три порядка за десять лет (50 GFLOPS/Вт);
- эффективное решение задач с интенсивной нерегулярной работой с памятью, в частности, задач на динамических графах, включая задачи принятия решений;
- эффективная обработка мощных потоков информации от источников разного типа;
- повышение надежности и отказоустойчивости по отношению к сбоям, отказам и информационным атакам;
- обеспечение продуктивности параллельного программирования в условиях резко выросшего уровня параллелизма (до 10^9 параллельных процессов).

Рост производительности на три порядка в системах экзафлопсного уровня планируется осуществить к 2018–2020 годам.

Традиционное разделение работ на эволюционные и инновационные дополнилось сегодня новым, третьим направлением, что произошло благодаря удивительно быстрому проникновению некоторых инновационных решений в коммерчески доступные компоненты. Это эмуляция инновационных архитектурно-программных принципов на эволюционно развивающихся кластерных суперкомпьютерах. Такие работы важны не только для отработки новых принципов — они позволяют более эффективно использовать эволюционирующие кластерные суперкомпьютеры. По-видимому, это крайне интересное и перспективное для российских разработчиков направление.

Современные метрики оценки производительности суперкомпьютерных систем

Кузьминский М.Б. (kus@free.net) — ФГБУН «Институт органической химии им. Н.Д. Зелинского РАН» (Москва)

В докладе приведен обзор современного состояния и перспектив применения тестов производительности высокопроизводительных систем. Рассмотрена возможная классификация тестов производительности. Приводятся примеры тестов прошлого и настоящего, а также тестов оценки будущих систем экзафлопсного уровня, для некоторых суперкомпьютеров дается оценка их преимуществ и недостатков. Обсуждаются причины, способствующие или препятствующие широкому распространению тестов на практике.

В докладе также анализируются некоторые результаты наиболее распространенных тестов, позволяющие сравнить применяемые вычислительные системы между собой и сделать выводы о тенденциях в развитии аппаратных средств, а также об эффективности применения тех или иных аппаратных и программных средств. В общем случае тесты позволяют дать количественную меру определенных качеств, которые включают производительность, надежность (например, MTBF), цену и ее «производные» (производительность на единицу стоимости) и др. Среди тестов производительности часто выделяют: микротесты, в том числе тесты компонентов компьютера (тесты производительности памяти — stream, ApexMap и др.), тесты подсистемы ввода/вывода (SPIOBENCH, IOR) или межсоединения. Для оценки производительности GPU используются, например, тесты SHOC и CLBenchmark.

Имеются также тесты типовых алгоритмов (Linpack, умножение матриц). Ядра (kernels) — типичные ключевые участки кодов, например ливерморские циклы. Синтетические тесты— например, тесты SPECcpu, SPECmpi и SPEComp, включающие части кодов из различных реальных приложений. Параллельные тесты, предполагающие распараллеливание на несколько процессоров или процессорных ядер при выполнении. Все другие представленные нами классы тестов также могут быть параллельными.

Для оценки производительности суперкомпьютерных систем предлагается оценивать ряд важных качеств HPC-систем, таких как цена, потребляемая электроэнергия, занимаемые площади и др., которые могут использоваться для расчета «производных» метрик производительности, например GFLOPS на доллар и т. п. Наиболее целесообразно делить производительность на полную стоимость владения компьютерной конфигурацией, однако для этого требуется договориться о едином способе ее подсчета.

Наиболее распространенными тестами производительности процессоров сегодня являются SpecCPU2006, и лидирующими по производительности процессорами на этих тестах на данный момент являются Intel Xeon Sandy Bridge, опережающие IBM Power7 и AMD Opteron. Недостатком SPECcpu2006 является разрешение применять автораспараллеливание компиляторами. Эффект от автораспараллеливания на используемых программах небольшой, поскольку коды не были предназначены для распараллеливания. Но и пренебречь влиянием автораспараллеливания на результаты нельзя, особенно если при сопоставлении разных процессоров результаты близки. Стремление производителей показать более высокий результат привело к тому, что данные в последовательном режиме стали публиковаться редко, что приводит к путанице в оценке возможностей отдельных процессорных ядер (последовательного выполнения). Но выполнять программы, подобные включенным в эти тесты, в режиме с автораспараллеливанием невыгодно — оно малоэффективно, для повышения пропускной способности вычислительной системы лучше одновременно выполнять последовательные программы. Для оценки возможностей распараллеливания некоторую информацию дают тесты SPECrate из состава SPECcpu, поскольку они показывают, не мешают ли несколько выполняющихся последовательных процессов друг другу из-за возможной конкуренции за ресурсы процессора.

Самым распространенным для оценки высокопроизводительных систем остается тест High Performance Linpack (HPL), на основании результатов которого строится рейтинг Top500. HPL может легко масштабироваться по размерности и служить тестом и для петафлопсных, и для будущих экзафлопсных систем. HPL критиковали за ограничение одной лишь задачей решения системы линейных уравнений, поэтому, в рамках набора тестов HPC Challenge (HPCC), он был дополнен тестами параллельного транспонирования матриц, быстрого преобразования Фурье, тестами памяти со случайным доступом, тестами задержек и пропускной способности коммуникаций. Другим примером возможного теста экзафлопсных систем является NPB (NAS Parallel Benhcmark).

На данный момент лидером Top500 является система Sequoia на базе IBM BlueGene/Q с показателем 16 PFLOPS, на втором месте – компьютер K от Fujitsu с производительностью свыше 10 PFLOPS, а все компьютеры в первой двадцатке Top500 превзошли порог 1 PFLOPS.

У других членов первой двадцатки количество процессорных ядер достигает ста тысяч,

поэтому встает вопрос о приложениях, которые эффективно распараллеливаются как для работы на таких петафлопсных системах так и на будущих экзафлопсных системах и могут применяться в качестве тестов. К числу таких приложений можно отнести некоторые задачи гидро/аэродинамики, вычислительной химии, квантовой хромодинамики и др.

Вычислительная химия использует суперкомпьютерные ресурсы для решения задач молекулярной динамики и квантовой химии. Распараллеливание молекулярной динамики обычно тормозится задержками межсоединения, однако в программном продукте NAMD достигнуто хорошее распараллеливание, он мало восприимчив к задержкам и был предложен как возможный тест петафлопсных систем. Он хорошо распараллеливается на 8 тыс. ядер, а далее узким местом может оказаться межсоединение.

Для хорошего распараллеливания типичных методов квантовой химии на очень большое число процессоров сегодня требуется введение приближений. Так, метод MP2 хорошо масштабируется в кластерах до 240 ядер, а для эффективной работы с 86 тыс. ядер в К-компьютере уже требуется введение приближения FMO (фрагментарных молекулярных орбиталей). Самый большой расчет по FMO на BlueGene/Q использовал 131 тыс. ядер. Метод DFT на специализированной для больших систем программе OPENMX хорошо масштабируется до 320 ядер, а его модификация хорошо распараллеливается и на 32 тыс. ядер К-компьютера, но при 131 тыс. ядер эффективность падает.

Другой прогнозируемый компонент экзафлопсных систем, присутствующий и в петафлопсных системах, — это GPU, применение которых, например, для больших квантово-химических расчетов связано с умножением больших плотных матриц. Вероятно, для систем с разреженными матрицами эффект будет существенно ниже. Другая область проведения оценок — расчеты методом конечных разностей. Известная программа Fluent включает тест обтекания кузова автомобиля с 14 млн ячеек. В Infiniband-кластерах, содержащих 512 ядер Intel Xeon E5, достигается эффективность распараллеливания на уровне 70%, а на 1024 ядрах она уже близка к 50%. Все это показывает, что нужно очень тщательно выбирать приложения для тестов петафлопсных и экзафлопсных систем.

Для измерения производительности HPC-систем в расчете на 1 Вт предложен, например, тест GBench. Кроме рейтинга Top500, применяется также и список Green500, в котором производительность на HPL делится на энергопотребление. Лидерами в производительности на ватт, как и в Top500, здесь пока являются IBM BlueGene/Q с показателем 2,1 GFLOPS/Вт и кластер на базе Intel Xeon E5 и Infiniband FDR с показателем 1,4 GFLOPS/Вт.

Не может существовать один идеальный тест для HPC. Тесты для HPC развиваются и используют возрастающие размерности по мере роста производительности суперкомпьютеров. Необходим комплексный подход, включающий применение различных типов тестов — от микротестов отдельных компонентов аппаратуры и тестов средств разработки до тестов реальных приложений. Синтетический тест может иметь результатом одну усредненную величину, однако его целесообразно дополнять данными других тестов. Целесообразно получать в рамках теста сразу несколько результатов подобно тому, как это делается в HPCC. Тесты обязательно должны включать как последовательное, так и параллельное выполнение. Однако в связи с повсеместным использованием многоядерных процессоров некоторые тесты производительности стали выполняться чуть ли не исключительно в распараллеленном виде, что затрудняет аккуратный сопоставительный анализ.

К числу наиболее популярных сегодня тестов, которые следует использовать и для оценки высокопроизводительных систем, следует отнести stream, SPECcpu2006, HPL, HPCC, SpecOMP2001, SpecMPI2007 и др. Однако следует учесть, что распространенность тестов на практике и соответствующая репрезентативность их результатов определяются не только научно-практической значимостью теста, но и организацией сбора результатов, соображениями маркетинга и другими подобными факторами.

Эволюция микропроцессорных архитектур

Корнеев В.В. (korv@rdi-kvant.ru) — ФГУП «НИИ “Квант”» (Москва)

В период между появлением микропроцессора Intel 80860 и кристалла векторного процессора японского суперкомпьютера Earth Simulator понятия «процессор» и «микропроцессор» слились — все, что может придумать архитектор процессоров, возможно сделать на одном кристалле. Анализ ретроспективы развития микроэлектроники показывает, что как только появляется технологическая возможность, в кристалле реализуется вся функциональность материнских плат и блоков, состоящих из совокупности плат, за исключением, пожалуй, памяти большого объема. Поэтому с большой уверенностью представляется возможным прогнозировать развитие микропроцессорных архитектур, следуя эволюционному пути совершенствования микропроцессорных кристаллов и архитектур суперкомпьютеров.

Следует, однако, учесть, что сети многопроцессорных систем, формируемых из однокристальных процессоров, и сети процессорных ядер, размещенных на одном кристалле, имеют разный набор ограничений пропускной способности [1]. В первом случае в качестве ограничений выступают число выводов кристалла, лимитирующих ширину линий, а также энергетические затраты, требуемые приемо-передатчиками и линиями связи. В последнем случае ограничений на ширину линий нет, равно как нет больших энергетических затрат в линиях, однако возникают технологические ограничения разводки линий, особенно широких линий, по кристаллу. Поэтому в накристальных сетях важным представляется допустимое число уровней металлизации для разводки проводников в разных слоях и выбор топологии межпроцессорных связей, зависящий также от количества объединяемых процессоров.

Коммуникационный кристалл YARC [2] демонстрирует возможность создания кристаллов с большим числом портов высокой пропускной способности. Нет сомнений, что в существенно многоядерный кристалл могут быть внедрены сериализаторы-десериализаторы, обеспечивающие высокую бодовую скорость при малой разрядности портов (выводов кристалла, образующих канал). Узлы с такого рода коммуникационными возможностями могут быть объединены в суперкомпьютеры с различными топологиями для создания сетей с требуемой проблемной ориентацией: многомерные кубы, булевские гиперкубы высокой размерности, сложенные сети Клоза и др.

На сегодня в суперкомпьютерах применяются многоядерные микропроцессорные кристаллы со встроенными контроллерами памяти и каналами «точка-точка», такими как HyperTransport AMD и QuickPath Intel. Это увеличивает производительность и коммуникационные возможности по доступу к локальной и удаленной памяти.

Для обеспечения высокой пропускной способности встроенные в кристалл контроллеры памяти должны поддерживать высокую степень расслоения локального блока памяти, а процессорные ядра – допускать большое число незавершенных обращений к памяти. На уровне ОС суперкомпьютера должно программно формироваться глобальное адресное пространство разделяемой памяти, состоящей из блоков локальных памятей узлов.

Следует ожидать, что в ядра многоядерных кристаллов будут введены векторная обработка и мультитредовость, а также элементы программируемой логики, наряду с уже присутствующей сейчас SIMD-обработкой. Иными словами, функциональность ядер приобретет функциональность узлов сегодняшних суперкомпьютеров, включая узлы с одним или несколькими графическими процессорами GP GPU.

Существенно многоядерные кристаллы с 64 ядрами и более могут быть использованы как в качестве универсальных процессоров суперкомпьютеров, так и как сопроцессоры, ориентированные на реализацию потоковых вычислений. Например, кристаллы Tile 64 и MIC (Xeon Phi) могут быть применены как в том, так и в другом качестве.

Асинхронные и синхронные треды. Одна и та же параллельная обработка данных в микропроцессорах может выполняться асинхронными и синхронными тредами. Каждый асинхронный тред имеет собственный счетчик команд, а параллелизм носит характер SPMD,

синхронные треды имеют один общий счетчик команд на всех, исполнение в стиле SIMD. Возможна иерархическая архитектура, в которой каждый асинхронный тред выполняет собственную совокупность синхронных тредов. Считается, что синхронные треды более сложны в программировании, но их использование позволяет увеличить производительность за счет меньших затрат аппаратуры на каждый тред и меньших затрат энергии.

Организация доступа в память. При исполнении программ каждое процессорное ядро узла суперкомпьютера обращается к распределенной разделяемой памяти. Так как дальнейшее выполнение программы возможно только после получения операнда из памяти, в случае чтения или подтверждения завершения записи перед выполнением последующего чтения, то в процессорных ядрах для сокращения простоев, связанных с ожиданием операндов или синхронизацией, необходимо вводить механизм, уменьшающий эти простои. Данная проблема связана с возрастающим разрывом в быстродействии логических элементов и элементов памяти, однако в параллельных системах она приобретает дополнительную остроту в связи с тем, что пропускная способность коммуникационной среды много ниже пропускной способности доступа к локальному блоку памяти.

Для преодоления негативного влияния на производительность времени доступа к памяти развивается подход, основанный на поддержке протекания «легких» тредов [3–5]. Загрузка процессора возложена на компилятор и программиста, которые должны выявлять в программе как можно больше даже очень небольших последовательностей команд, оформляемых как легкие треды и асинхронно запускаемых на исполнение при наличии простаивающих ресурсов процессора. Этот подход к заданию синхронизации легких тредов служит альтернативой используемому при синхронизации POSIX-тредов. Широко используемый подход к реализации синхронизации на базе «замков» (переменных, чаще всего битов, значения которых проверяются и изменяются неделимыми последовательностями команд) представляется малопродуктивным при синхронизации динамически порождаемых тредов, в том числе легких тредов, из-за больших задержек. Поэтому предлагается синхронизация на базе расширения каждого слова памяти дополнительным FE-битом full/empty, значение которого full устанавливает, что слово памяти имеет содержимое, — в противовес отсутствию содержимого при значении empty. Это позволит выдавать максимально возможное в исполняемой программе количество обращений к памяти, генерируемых легкими тредами, и параллельно выполнять эти доступы в расслоенной памяти.

Потоковая обработка. Необходимо минимизировать количество обращений к памяти, что возможно, например, при выборе алгоритма решения задачи, допускающего создание потоковой программы. В этом случае прикладная программа подготавливается пользователем исходно как потоковая или преобразуется в таковую компилятором. Межъядерные потоки программируются исходя из параметрического описания графов связей подсистем, на которых эти программы способны выполняться [6]. Для выполнения прикладных программ операционная система суперкомпьютера должна формировать связную подсистему процессорных ядер с требуемым программой типом графа межъядерных связей и алгоритмом нумерации ядер. При потоковом представлении программы значительная доля обрабатываемых операндов будет передаваться между ядрами по линиям, существенно снижая количество обращений в память. Сегодня имеются примеры создания потоковых программ для ряда задач линейной алгебры и математической физики с ускорением параллельных вычислений, прямо пропорциональным количеству используемых процессорных ядер при соответствующем размере задачи.

Литература

1. D. Wentzlaff et al. On-Chip Interconnection Architecture of the Tile Processor. IEEE Micro. September-October. 2007. pp. 15–31.

2. S. Scott et al. The BlackWidow High-Radix Clos Network.Proceedings of the 33rd annual international symposium on Computer Architecture. 2006. pp. 16–28.

3. K. Wheeler et al. Qthreads: An API for Programming with Millions of Lightweight Threads.

Proceedings of 22nd IEEE International Parallel and Distributed Processing Symposium. IPDPS 2008.
4. S. Li et al. A Heterogeneous Lightweight Multithreaded Architecture. IEEE International Parallel and Distributed Processing Symposium, 2007.
5. P. Kogge et al. Computer Systems with Lightweight Multi-Threaded Architectures. PatentUS 2007/0198785 A1, August 23, 2007.
6. Корнеев В. В. Архитектура вычислительных систем с программируемой структурой. Новосибирск: Наука, 1985.

СЕКЦИЯ.
ЭКЗАФЛОПСНЫЕ ТЕХНОЛОГИИ

Опыт разработки отечественной высокоскоростной коммуникационной сети для суперкомпьютеров

Симонов А.С., Слуцкин А.И., Макагон Д.В., Сыромятников Е.Л., Жабин И.А., Фролов А.С., Щербак А.Н. ({simonov, slutskin, makagond, syromiatnikov, zhabin, frolov, andrey.shcherbak}@nicevt.ru) — ОАО «НИЦЭВТ» (Москва)

Производительность суперкомпьютеров на современных задачах в значительной степени зависит от свойств высокоскоростной межузловой коммуникационной сети. Коммерчески доступные сети в силу своей универсальности и, как следствие, существенной избыточности не позволяют получать приемлемые значения основных характеристик, определяющих эффективность вычислительной системы при решении реальных вычислительно сложных задач. В ОАО «НИЦЭВТ» выполняется разработка высокоскоростной отказоустойчивой коммуникационной сети «Ангара» с топологией 4D-тор, которая может стать основой для создания отечественных технологий разработки суперкомпьютеров, при этом обеспечивается возможность объединения десятков тысяч вычислительных узлов.

При разработке концепции коммуникационной сети был принят ряд решений относительно модели исполнения прикладных задач на вычислительной системе. Одно из наиболее важных состоит в том, что на вычислительном узле исполняются процессы только одной прикладной задачи. Это позволяет, во-первых, сократить накладные расходы на переключение между процессами, а во-вторых, означает поддержку только двух виртуальных адресных пространств: решаемой задачи и операционной системы. Последняя, помимо своих основных функций, необходима и для реализации параллельной файловой системы, функционирующей поверх коммуникационной сети и обеспечивающей возможность выполнения не только стандартных операций ввода/вывода, но и реализации контрольных точек.

Для объединения узлов суперкомпьютерной системы используется топология 4D-тор. При этом в каждый вычислительный узел или узел ввода/вывода вычислительной системы устанавливаются маршрутизатор и специализированное программное обеспечение, обеспечивающие поддержку распределенной глобально адресуемой памяти в рамках прикладной задачи. Выбор топологии «многомерный тор» сделан исходя из особенностей шаблонов обмена сообщениями по сети на интересующем классе задач [1,2]. По сравнению с топологией Fat tree, используемой в сети Infiniband, данная топология более толерантна к неравномерной загрузке, что подтверждается не только результатами имитационного моделирования, но и тестовыми расчетами на существующих суперкомпьютерах, например IBM Blue Gene/P с сетью топологии 3D-тор [2]. Кроме того, обеспечивается хорошая масштабируемость производительности вычислительной системы при решении задач с интенсивным обменом сообщениями.

Разработка данной коммуникационной сети ведется в ОАО «НИЦЭВТ» с 2006 года [3,4]. На подготовительном этапе было проведено тщательное изучение результатов зарубежных исследований и разработок, в том числе работ Уильяма Дэйли [5] и Хосе Дуато [6], а также архитектур IBM Blue Gene [7] и Cray SeaStar/Gemini [8]. За это время были созданы три поколения макетных образцов маршрутизаторов на FPGA и вычислительные кластеры на их основе, отработаны решения по передаче сообщений и взаимодействию с коммерчески доступными процессорами вычислительных узлов с использованием современного интерфейса PCI Express.

В настоящее время завершена разработка заказной СБИС маршрутизатора ЕС8430 на технологических нормах 65 нм, данные для ее изготовления переданы на зарубежную фабрику. Отлажен комплект системного программного обеспечения, включающий драйвер маршрутизатора, создана пользовательская библиотека нижнего уровня, реализована поддержка основных стандартов параллельного программирования: CRAY SHMEM, MPI 2.0, OpenMP, UPC, GASNet, ARMCI, стандартных математических библиотек BLAS, LAPACK, SCALAPACK, FFTW и др. Для окончательной отладки программного обеспечения маршрутизаторов ЕС8430

на основе СБИС завершаются работы по изготовлению прототипного кластера, состоящего из 36 узлов. Структурная схема маршрутизатора ЕС8430 «Ангара» приведена на рисунке.

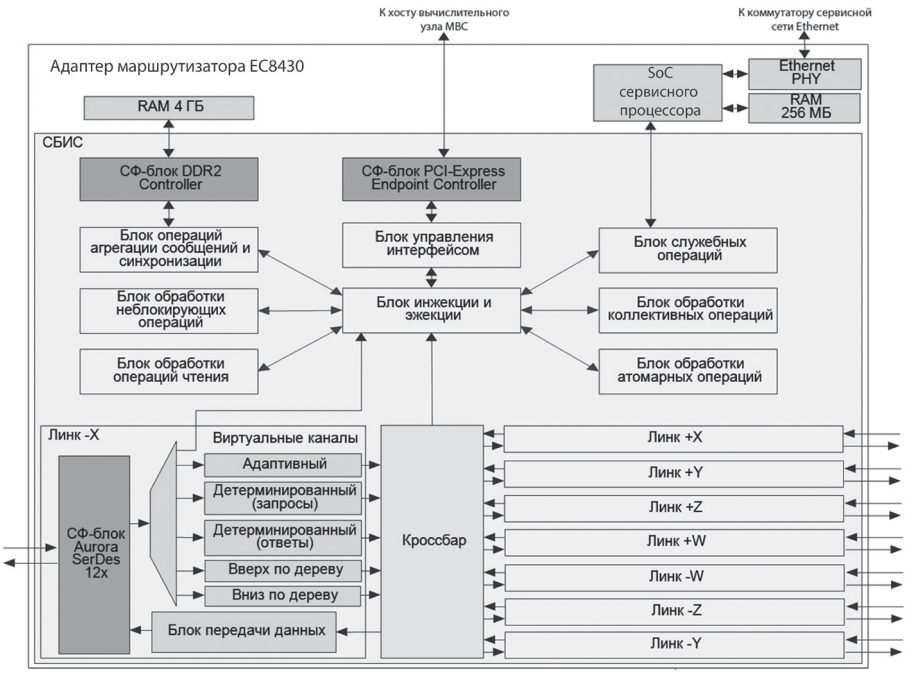

Рисунок. Структурная схема маршрутизатора ЕС8430 «Ангара»

В маршрутизаторе ЕС8430 «Ангара» реализованы следующие функциональные возможности: протокол надежной передачи пакетов на канальном уровне; детерминированная и адаптивная передача данных; аппаратная поддержка многопоточности; аппаратная реализация операций (запись в память удаленного узла; запись в память удаленного узла с сохранением концептуальности; атомарные операции в памяти удаленного узла; чтение из памяти удаленного узла; неблокирующая запись больших массивов в память удаленного узла; чтение больших массивов из памяти удаленного узла); аппаратная поддержка операций барьерной синхронизации; сборка массивов в памяти маршрутизатора с последующим копированием в память узла; выполнение коллективных операций (операции broadcast, reduce); система обеспечения отказоустойчивости.

Основным рекомендуемым режимом работы разрабатываемой коммуникационной сети, на котором обеспечиваются наилучшие показатели производительности суперкомпьютера, является режим с использованием библиотеки SHMEM. При этом общий объем глобально адресуемой маршрутизатором памяти составляет до 128 ПБайт, а объем памяти, адресуемой на каждом узле, — до 2 ТБайт.

По функциональным возможностям и производительности сеть «Ангара» соответствует мировому уровню. Основные характеристики сети «Ангара», коммуникационная задержка и пропускная способность, сопоставимы с лучшими образцами заказных коммуникационных сетей зарубежных суперкомпьютеров компаний CRAY, IBM, Fujitsu, занимающих верхние строчки списка Top500 (см. таблицу).

Таблица. Сравнение характеристик коммуникационной сети ЕС8430 с коммерчески доступными и заказными коммуникационными сетями

Характеристика	Infiniband 4x FDR	Cray Gemini	Макет на ПЛИС	СБИС ЕС8430
Интерфейс с процессором вычислительного узла	PCI Express x8 gen 3	HyperTransport 3.0	PCI Express x8 gen 1	PCI Express x16 gen 2
Топология	Fat tree	3D-тор	2D-тор	4D-тор
Пропускная способность канала связи, Гбайт/с	7,25	9,375	1,25	9,375
Задержка точка-точка (соседние узлы), мкс	1,0	1,4	2,1	< 1,0
Задержка на хоп, мкс	0,165–0,495	0,1	0,4	< 0,2
Масштабируемость	48K/384K	100K	256	32K

По сравнению с коммуникационными сетями с топологией 3D-тор, использование коммуникационной сети с топологией 4D-тор имеет ряд преимуществ, таких как повышенная связность и бисекционная пропускная способность, повышенная отказоустойчивость, естественная применимость на более широком спектре физических задач и задач моделирования.

Продвижение сети «Ангара» на рынок планируется осуществлять в трех вариантах: в виде СБИС для использования в составе серверных платформ сторонних производителей, как отдельной коммерческой сети в виде адаптеров PCI Express для кластерных систем со стандартными процессорами и чипсетами и в составе разрабатываемой в ОАО «НИЦЭВТ» вычислительной платформы «Ангара» в форм-факторе blade.

Литература

1. Фролов А.С., Семенов А.С., Мошкин Д.В., Кабыкин В.К., Никитин А.И. Суперкомпьютеры для графовых задач // Открытые системы. СУБД. 2011. № 7

2. Пожилов И.А., Семенов А.С., Макагон Д.В. Прогнозирование масштабируемости задачи умножения разреженной матрицы на вектор при помощи модели коммуникационной сети // Материалы конференции: Научный сервис в сети Интернет. Абрау-Дюрсо, 19–24 сентября 2011.

3. Слуцкин А.И., Симонов А.С. Развитие суперкомпьютерных технологий в ОАО «НИЦЭВТ» // Научно-техническая конференция: Перспективные направления развития средств вычислительной техники: Сборник тезисов докладов. Москва, 28 июня 2011.

4. Симонов А.С., Жабин И.А., Макагон Д.В. Разработка межузловой коммуникационной сети с топологией «многомерный тор» и поддержкой глобально адресуемой памяти для перспективных отечественных суперкомпьютеров // Научно-техническая конференция: Перспективные направления развития средств вычислительной техники: Сборник тезисов докладов. Москва, 28 июня 2011.

5. William Dally and Brian Towles. Principles and Practices of Interconnection Networks. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.

6. Jose Duato, Sudhakar Yalamanchili, and Ni Lionel. Interconnection Networks: An Engineering Approach. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.

7. N.R. Adiga, M.A. Blumrich, D. Chen, P. Coteus, A. Gara, M.E.Giampapa, P.Heidelberger, S. Singh, B.D. Steinmacher-Burow, T. Takken, M. Tsao, P. Vranas. Blue Gene/L torus interconnection network, IBM J. RES. & DEV. VOL. 49 NO. 2/3 MARCH/MAY 2005.

8. Robert Alverson, Duncan Roweth, and Larry Kaplan. The Gemini System Interconnect, In Proceedings of the 2010 18th IEEE Symposium on High Performance Interconnects (HOTI '10), IEEE Computer Society, Washington, DC, USA, pp. 83–87.

Перспективы виртуализации суперкомпьютерных систем

Кудрявцев А.О. (alexk@ispras.ru), Кошелев В.К. (vedun@ispras.ru), Аветисян А.И. (arut@ispras.ru) — ИСП РАН (Москва)

В настоящее время технологии виртуализации серверов широко применяются в индустрии и для типичных задач обеспечивается достаточный уровень производительности. Развитие технологий привело к возможности использования виртуализации в области высокопроизводительных вычислений, однако возникающие накладные расходы зачастую делают такое применение нецелесообразным. В докладе анализируются перспективы применения технологий виртуализации в области высокопроизводительных вычислений на платформе x86-64. Рассматриваются основные причины падения производительности при запуске параллельных программ в виртуальной среде. Рассматриваются конкретные системы виртуализации KVM/QEMU и Palacios, в качестве тестовых пакетов используются HPC Challenge и NAS Parallel Benchmarks. Тестирование выполняется на вычислительном кластере, построенном на базе высокоскоростной сети Infiniband.

Технологии аппаратной и контейнерной виртуализации позволяют сегодня обеспечить производительность отдельных классов приложений в виртуальных средах, практически неотличимую от производительности на реальном оборудовании. Исследователи по всему миру начали изучать возможности и ограничения виртуализации при использовании в области высокопроизводительных вычислений (HPC, High Performance Computing). Преимущества виртуализации в области высокопроизводительных вычислений широко обсуждаются [1,2]: устойчивость к сбоям, совместимость и гибкость при работе с виртуальными машинами (ВМ). Также весьма интересна идея применения концепции облаков для создания высокопроизводительных масштабируемых систем, позволяющих эффективно использовать ресурсы. Облачные системы такого типа позволят организовывать доступ к высокопроизводительным системам как к сервису, например, сервис «виртуальный MPI-кластер» позволит пользователю по запросу создавать MPI-кластер с требуемыми характеристиками.

Исследования показывают, что виртуализация HPC имеет смысл по крайней мере для отдельных классов приложений [1], однако, несмотря на это, сегодня существует серьезная нехватка экспериментальных данных. Необходимо исследовать поведение различных приложений при запуске на различном оборудовании с использованием ряда систем виртуализации, для того чтобы достоверно оценить ограничения и возможности существующих технологий. Современные многопроцессорные, многоядерные системы предъявляют новые требования к системам виртуализации, включая корректную эмуляцию архитектуры NUMA (Non-Uniform Memory Access, архитектура с неравномерным доступом к памяти) в гостевой системе.

Основной целью работы, которой посвящен доклад, является оценка накладных расходов виртуализации в целом, а также выявление причин и минимизация этих расходов. Одной из задач было достижение максимально возможной производительности при выполнении заданий, запущенных в виртуальной среде, по сравнению с запуском на реальном оборудовании. Исследования проводились с использованием системы виртуализации KVM (Kernel-based Virtual Machine) [3] и гипервизора Palacios [4], который изначально разрабатывался специально для высокопроизводительных систем. Необходимо отметить, что в ходе работы использовалась специально модифицированная версия гипервизора Palacios, поскольку оригинальная не может быть запущена на используемом тестовом оборудовании.

Для достижения наилучшей производительности (относительно случая запуска на реальном оборудовании) виртуальным машинам выделяется максимально возможное количество ресурсов, включая все процессорные ядра. Помимо этого, виртуальным машинам предоставляется реальное коммуникационное устройство, с использованием технологий виртуализации ввода/вывода Intel VT-d в случае KVM и паравиртуализации в случае Palacios. Также в ВМ эмулируется архитектура NUMA в соответствии с реальной конфигурацией узлов кластера.

В качестве тестовых пакетов используются пакеты HPC Challenge [5] и NAS Parallel Benchmarks [6]. Тестирование производительности выполнялось на современном кластере, с использованием до 8 узлов, построенных на базе процессоров Intel Xeon (в сумме до 96 процессорных ядер). Узлы кластера связаны сетью Infiniband 40 Гбит/с. Стоит отметить, что многие ранние исследования в области виртуализации HPC проводились на системах из одного узла, что не позволяло полноценно оценить накладные расходы, вызываемые виртуализацией.

В ходе выполнения данной работы были достигнуты следующие результаты:

- Описаны базовые, эффективные методы настройки KVM/QEMU для запуска HPC-приложений в виртуальной среде. Показана необходимость корректной эмуляции архитектуры NUMA в виртуальной среде в соответствии с реальной топологией системы.
- Проведена оценка накладных расходов виртуализации для современного HPC-кластера, построенного на базе Infiniband. Узлы кластера — двухпроцессорные 12-ядерные серверы архитектуры NUMA, используется до 8 узлов.
- Выполнено сравнение гипервизора Palacios, встроенного в основную ОС Kitten, и системы виртуализации KVM/QEMU, с основной ОС Linux. Полученные данные показывают, что KVM/QEMU дает более стабильные и предсказуемые результаты, в то время как Palacios имеет меньшие накладные расходы на «мелкозернистых» тестах, особенно при использовании большого числа вычислительных ядер.
- Проведено исследование полученных тестовых данных и накладных расходов, вызванных системой виртуализации. В частности, исследовано влияние гранулярности коммуникаций HPC-приложения и частоты прерываний на величину накладных расходов.

Литература

1. A.J. Younge, R. Henschel, J. Brown, G. von Laszewski, J. Qiu, and G. C. Fox. Analysis of Virtualization Technologies for High Performance Computing Environments. In The 4th International Conference on Cloud Computing (IEEE CLOUD 2011), July 2011.

2. A. Gavrilovska, S. Kumar, H. Raj, K. Schwan, V. Gupta, R. Nathuji, R. Niranjan, A. Ranadive, and P. Saraiya. Abstract High-Performance Hypervisor Architectures: Virtualization in HPC Systems. In 1st Workshop on System-level Virtualization for High Performance Computing (HPCVirt), in conjunction with EuroSys 2007, 2007.

3. A. Kivity, Y. Kamay, D. Laor, U. Lublin, and A. Liguori. KVM: the Linux virtual machine monitor. In OLS '07: The 2007 Ottawa Linux Symposium, pp. 225–230, July 2007.

4. J. R. Lange, K. Pedretti, P. Dinda, P. G. Bridges, C. Bae, P. Soltero, and A. Merritt. Minimal-overhead virtualization of a large scale supercomputer. In Proceedings of the 7th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments, VEE '11, pp. 169–180, 2011.

5. P. R. Luszczek, D. H. Bailey, J. J. Dongarra, J. Kepner, R. F. Lucas, R. Rabenseifner, and D. Takahashi. The HPC Challenge (HPCC) benchmark suite. In Proceedings of the 2006 ACM/IEEE conference on Supercomputing, SC '06, 2006.

6.D. Bailey, T. Harris, W. Saphir, R. van der Wijngaart, A. Woo, and M. Yarrow. The NAS parallel benchmarks 2.0. Technical Report NAS-95-020, NASA Ames Research Center. December 1995.

Архитектура «РСК Торнадо»: преимущества и энергоэффективность

Московский А.А. (moskov@rsc-tech.ru) — ЗАО «РСК Технологии» (Москва)

Рост энергопотребления современных ЦОД выдвинул на передний план такой показатель, как энергоэффективность, который стал ключевым параметром оценки решений не только центров обработки данных, но и суперкомпьютеров. Такая метрика, как коэффициент полезного использования мощности или эффективности использования электроэнергии (PUE, Power Usage Effectiveness), вычисляемая как отношение общего энергопотребления ЦОД к потреблению только ИТ-оборудования, в большинстве современных ЦОД оказывается в промежутке от 1,5 до 2, что говорит о резерве для оптимизации расходов энергии на поддержку инфраструктуры – в первую очередь на подсистему охлаждения.

Жидкостное охлаждение как передовая технология, позволяющая радикально повысить энергоэффективность суперкомпьютеров, становится все более актуальным и широко востребованным в индустрии. За последние два года количество решений с жидкостным охлаждением выросло на порядок, хотя еще три года назад такие решения были редкостью, а сегодня развиваются альтернативные подходы к методам жидкостного охлаждения, хотя их эффективность в большинстве случаев не была доказана на практике.

В докладе иллюстрируется эффективность технологии прямого, контактного жидкостного охлаждения с использованием охлаждающих пластин-радиаторов (coldplates) на примере архитектуры «РСК Торнадо», разработанной и реализованной в ряде крупных проектов в России. Проведено сравнение затрат энергии на выполнение реального расчета с использованием одной и той же аппаратной платформы, с использованием охлаждения принудительным обдувом воздухом и с жидкостным охлаждением (архитектура «РСК Торнадо»). Показано, что общая экономия электроэнергии с учетом сниженного энергопотребления процессора (за счет более низкой рабочей температуры) и инфраструктурных факторов составляет около 80%. Следует отметить, что архитектура «РСК Торнадо» построена на базе стандартных серверных компонентов: массово доступных серверных плат (выпущенных различными производителями и изначально созданных для традиционных систем с воздушным обдувом электронных компонентов), процессоров Intel Xeon (включая модели Intel Xeon E5-2690), причем обеспечивается возможность постоянной работы технологии Intel Turbo Boost, что дает прирост тактовой частоты на 400 МГц выше номинальной. Все это позволяет сочетать доступность по цене и срокам с такими преимуществами решений высокого уровня, как надежность и высокая эффективность.

Архитектура «РСК Торнадо» легко масштабируется. Решение РСК микроЦОД (от 16 до 64 вычислительных узлов) можно разместить в одном монтажном шкафу (менее 2 кв. м площади) со всей необходимой инфраструктурой (за исключением наружного модуля подсистемы охлаждения). Для системы РСК мини-ЦОД (до 256 узлов) необходимо лишь два изолированных монтажных шкафа (до 4 кв. м площади): один для размещения вычислительных узлов, а второй – вспомогательный – для сетевого оборудования, а также внутреннего блока подсистемы охлаждения. Небольшие установки «РСК Торнадо» обладают столь же высокой эффективностью, как и большие системы. Замеры PUE для системы класса РСК мини-ЦОД, установленной в Росгидромете (96 двухпроцессорных узлов на базе Intel Xeon E5-2690), продемонстрировали рекордный для индустрии показатель энергоэффективности на уровне 1,06 в режиме free-cooling.

Группа компаний РСК с 2009 года реализовала в России ряд крупных проектов по разработке и установке высокопроизводительных вычислительных систем с жидкостным охлаждением (например, в Южно-Уральском государственном университете, Московском физико-техническом институте и др.), что на сегодняшний день позволяет говорить о наличии у ее специалистов большого опыта реализации и эксплуатации суперкомпьютеров с такой технологией охлаждения. Например, суперкомпьютер класса РСК ЦОД, установленный в ЮУрГУ и прошедший плановую модернизацию, занял в 2011 году 86-ю позицию в рейтинге

Top500. Кроме того, эта система оказалась самой энергоэффективной в России по данным рейтинга Green500 в 2011–2012 гг. Во второй половине 2012 года было представлено третье поколение архитектуры «РСК Торнадо» с использованием новейших сопроцессоров Intel Xeon Phi.

Решения IBM в области высокопроизводительных вычислений

Горбас С.А. (sgorbas@ru.ibm.com) — IBM, Россия и СНГ (Москва)

Автоматическое отображение высокоуровневых программ на современные параллельные вычислительные системы со сложной архитектурой

Штейнберг Б.Я. (borsteinb@mail.ru) — Южный федеральный университет (Ростов-на-Дону)

Потребность в быстрых вычислениях возникает не только в известных приложениях суперкомпьютеров, входящих в список Top500, но и, например, в задачах искусственного интеллекта, решаемых роботами (которые не могут быть оснащены суперкомпьютером), — в таких случаях нужны высокопроизводительные процессоры или небольшие ускорители (платы).

В 2006 году ведущие производители микроэлектроники начали выпуск многоядерных процессоров, а параллельные вычисления приобрели массовый характер: процессоры Intel, AMD Opteron6000, 512-ядерный графический ускоритель NVidia-Fermi, 9-ядерный IBM Cell BE, 96-ядерный ускоритель ClearSpeed Advance X620, 64-ядерный процессор Tile64 и другие высокопроизводительные процессоры и платы, включая программируемые логические матрицы (ПЛИС), используемые в качестве ускорителей. Перечисленные архитектуры отличаются не только количеством ядер, но и способом их соединений, видами параллельных вычислений (SIMD, MIMD, Pipeline), используемой памятью и другими особенностями. Каждая высокопроизводительная архитектура эффективна на некотором своем классе прикладных задач. Наблюдается тенденция расширения многообразия вычислительных архитектур. Далеко не все архитектуры поддерживают систему команд x86. Все это приводит к проблеме создания программного обеспечения, адекватно использующего возможности новых вычислительных архитектур.

Решение проблемы быстрой и недорогой разработки эффективного программного обеспечения для новых вычислительных архитектур невозможно без новых инструментов разработки программ. Одним из основных таких инструментов является оптимизирующий распараллеливающий компилятор.

Современное поколение вычислительных систем, несмотря на их многообразие, характерно тем, что время передачи данных с кристалла оперативной памяти на кристалл процессора во много раз больше времени обработки этих данных. Многоядерность процессора имеет эффект для тех задач, в которых с малым числом данных выполняется много операций.

Площадь на кристалле процессора распределяется между памятью (кэш) и вычислительными ядрами — увеличение объема памяти и количества вычислительных ядер могут повышать быстродействие процессора. Чем меньшую площадь кристалла будут занимать вычислительные ядра, тем больше может быть кэш-память. Уже появляются многоядерные процессоры с количеством ядер порядка 100, но такое увеличение приводит к уменьшению памяти на этом же кристалле — ускорятся ли при этом вычисления? Оказывается, что для разных алгоритмов по-разному:

- алгоритмы, для которых увеличение кэш-памяти и увеличение количества ядер не приводят к увеличению быстродействия (скалярное произведение векторов) — время вычисления сводится ко времени перекачки данных с кристалла памяти на кристалл процессора;
- алгоритмы, для которых существенно увеличение кэш-памяти (сортировка);

- алгоритмы, для которых существенно увеличение количества ядер (вычисление степенных рядов);
- алгоритмы, быстродействие которых зависит от баланса между объемом кэш-памяти и количеством вычислительных ядер (перемножение матриц).

Если процессор с программируемой архитектурой допускает только изменение связей между процессорами на кристалле, то его эффективность все равно не выйдет за пределы класса алгоритмов, который определяется отношением площади кристалла, отведенной для памяти, к площади кристалла, отведенной для вычислительных ядер.

Работа компиляторов с памятью. Существует семейство компиляторов с языка программирования высокого уровня в язык описания электронных схем. Такого типа компилятор может стать частью компилятора для программируемых архитектур. Компилятор может взять на себя функции некоторого преобразования кода и такого размещения данных в памяти, при котором минимизируются обмены данными между кристаллами или между вычислительными элементами внутри кристалла.

Имеется много методов, направленных на преобразование алгоритмов к блочному виду (Tiling). Но можно пойти дальше — для многих блочных алгоритмов выгодно, чтобы матрицы (массивы) в оперативной памяти размещались по блокам, а не по столбцам (Фортран) или строкам (Си, Паскаль). Такая манипуляция приводит к сложной адресации и, несмотря на это, дает ускорение.

Для некоторых классов прикладных задач можно данные размещать в распределенной памяти с перекрытиями и получать при этом ускорение вычислений. При таком размещении происходит увеличение количества выполняемых арифметических операций, но уменьшается количество пересылок данных, что приводит к увеличению быстродействия. Следует отметить, что пересылки данных возникают не только между процессорами в высокопроизводительном кластере, но и между модулями локальной памяти процессорных элементов, расположенных на одном кристалле (типа Tile64).

Некоторая оптимизация обращений к памяти используется во многих современных компилирующих системах: LLVM, PGI, GCC, Rose compiler infrastructure (Lawrence Livermore National Laboratory).

Внутреннее представление компилирующих систем. Большинство известных систем (GCC, LLVM, Intel C++/Fortran Compiler, Microsoft Visual C++ и др.) автоматической оптимизации и распараллеливания программ имеют внутреннее представление низкого уровня, близкое к ассемблеру, ориентированное на генерацию команд x86 или близких к ним по уровню. Низкий уровень внутреннего представления удобен для создания оптимизирующих компиляторов с широкого класса языков, даже далеких от Си и Фортрана, — таких как Java. Но эти внутренние представления неудобны для генерации кода на вычислительные архитектуры, далекие от системы команд x86.

Высокоуровневое внутреннее представление имеется в распараллеливающих системах SUIF, «ДВОР» и новом проекте Rose Compiler. Распараллеливающие компиляторы PGI (Portland Group Incorporation) основаны на распараллеливающей системе SUIF, и, таким образом, используют высокоуровневое внутреннее представление. Высокоуровневая оптимизация может разрабатываться до создания ассемблера, что позволяет сократить сроки разработки распараллеливающих компиляторов.

Анализ информационных зависимостей. Несмотря на обилие методов определения информационных зависимостей, эти зависимости во многих случаях определяются неточно или долго. Видны следующие перспективы развития анализа информационных связей: использование решетчатых графов, распараллеливание долгих процедур анализа (например, анализа псевдонимов), использование диалоговой компиляции. Решетчатые графы для анализа и преобразования программ использует распараллеливающая система Pluto (библиотека PipLib), причем эта же система преобразует программы к блочному виду (tiling).

Диалоговый режим компиляции. Диалоговый режим может иметь преимущества по сравнению с автоматическим при уточнении информационных зависимостей и изменении

порядка выполнения операций. Соотношение ручной, диалоговой (полуавтоматической) и автоматической оптимизации (распараллеливания) выглядит следующим образом:

- время выполнения генерируемого кода: ручная компиляция < диалоговая компиляция < автоматическая компиляция;
- уровень квалификации программиста: ручная компиляция > диалоговая компиляция > автоматическая компиляция;
- время разработки программы: ручная компиляция > диалоговая компиляция > автоматическая компиляция;
- множество оптимизируемых и/или распараллеливаемых программ: ручная компиляция > диалоговая компиляция > автоматическая компиляция.

Диалоговый режим компиляции имеется у коммерческой распараллеливающей системы Parawise, имеющей закрытый код, на входе язык Фортран и ориентированной на распараллеливание для вычислительных систем с процессорами платформы x86.

При разработке высокопроизводительных программ оптимизирующие компиляторы могут:

- уменьшить сроки разработки;
- понизить требования к квалификации программистов и понизить стоимость работ;
- повысить надежность;
- упростить переносимость на уровне исходного высокоуровневого кода.

Высокоуровневые эффективные программы должны быть написаны в стиле, допускающем автоматический анализ, распараллеливание и локализацию кода и данных.

Представленная в докладе распараллеливающая система «ДВОР» содержит высокоуровневое внутреннее представление, блочно-аффинные распределения данных в распределенной памяти, решетчатые графы, автоматический расчет задержек в стартах конвейеров, конвертер в VHDL и позволяет создавать распараллеливающие компиляторы.

CLAVIRE: облачная платформа для высокопроизводительных вычислений

Бухановский А.В. (boukhanovsky@mail.ifmo.ru) — СПбГУ ИТМО (Санкт-Петербург)

Обеспечение доступности суперкомпьютерных систем для выполнения междисциплинарных научных исследований является одной из приоритетных задач «электронной науки» (eScience). В рамках облачной парадигмы этот процесс связан с необходимостью создания новых технологий, обеспечивающих не только виртуализацию и управление высокопроизводительными вычислительными ресурсами, но и поддержку экосистемы облачных сервисов и композитных приложений на их основе, функционирующих в неоднородной распределенной вычислительной среде. Данная возможность реализована в многофункциональной облачной инструментально-технологической платформе CLAVIRE (CLoud Applications VIRtual Environment), созданной в рамках комплексного проекта по реализации Постановления Правительства Российской Федерации от 9 апреля 2010 года № 218 «О мерах государственной поддержки развития кооперации российских высших учебных заведений и организаций, реализующих комплексные проекты по созданию высокотехнологичного производства».

Платформа CLAVIRE предназначена для создания, исполнения и предоставления сервисов доступа к предметно-ориентированным высокопроизводительным композитным приложениям, функционирующим в облаке неоднородных вычислительных ресурсов корпораций, центров компетенции, центров обработки данных, инфраструктур экстренных вычислений и распределенных хранилищ данных и знаний. На основе CLAVIRE возможно развертывание распределенного программно-аппаратного комплекса поддержки инфраструктуры предметно-ориентированных облачных вычислений в различных областях знания и технологий. Новизна и технологическая эффективность разработки определяется симбиозом применяемых для ее создания подходов: концепции облачных вычислений, интеллектуального управления распределенными вычислительными ресурсами и сервисными

технологиями, основанными на знаниях предметной области.
Основные преимущества CLAVIRE:

- универсальность —создание и поддержка облачных сред различного назначения;
- унификация —возможность подключения и использования разнородных вычислительных ресурсов и источников данных, включая серверы данных и приложений, суперкомпьютеры, грид, облака (IaaS и PaaS), сетевые хранилища данных, а также информационно-измерительные комплексы различного назначения;
- поддержка различных интерфейсов доступа к облачным ресурсам — возможность пользователям конструировать собственные композитные приложения, использовать готовые сервисы через проблемно-ориентированные веб-интерфейсы, а также через программные интерфейсы взаимодействия с локальными приложениями;
- поддержка технологий интерактивного управления облачными приложениями— возможность создания распределенных систем реального времени, систем интерактивной визуализации и виртуальной реальности, а также инфраструктуры ситуационных центров;
- информационная безопасность — наличие инновационной модели защиты от несанкционированного доступа к облачным ресурсам в системах с неопределенным контуром;
- интеллектуальная поддержка процессов расширения перечня ресурсов и сервисов за счет активности сообщества пользователей, включая автоматизацию установки и регистрации прикладного программного обеспечения в облаке;
- автоматическая генерация сложных графических веб-интерфейсов для прикладных пакетов и композитных приложений на основе их предметно-ориентированных описаний, вне зависимости от специфики среды исполнения;
- бесшовное и безопасное сопряжение облачных сервисов и композитных приложений с прикладным программным обеспечением на компьютере пользователя (технология V-Clouds);
- интеллектуальная поддержка создания и распространения проблемно-ориентированных коллекций прикладных сервисов на основе технологии виртуальных моделирующих объектов, позволяющей сделать процесс создания новых облачных приложений доступным для широкого круга специалистов-предметников;
- эффективное планирование исполнения пользовательских задач, включая автоматическое распараллеливание композитных приложений между отдельными физическими ресурсами в составе облачной среды и прогноз количества необходимых ресурсов;
- обеспечение прозрачной системы оценки стоимости и биллинга, исходя из фактического времени использования ресурсов и сервисов, с учетом интересов всех провайдеров внешних ресурсов и сервисов, входящих в композитное приложение;
- виртуальные лаборатории — возможность применения для целей обучения по различным предметным областям, с устранением эффектов эрозии знаний.

Платформа CLAVIRE обеспечивает весь жизненный цикл облачных сервисов в рамках моделей SaaS и AaaS, включая создание и исполнение предметно-ориентированных высокопроизводительных композитных приложений, функционирующих в неоднородной распределенной вычислительной среде, для различных задач науки, промышленности, бизнеса и социальной сферы. Платформа нашла успешное применение при реализации задач исследовательского проектирования морских судов и объектов океанотехники, предотвращения наводнений в Санкт-Петербурге, моделирования наноэлектронных устройств, прототипирования бортовых систем управления динамическими объектами, создания анимационных фильмов, а также анализа и управления информационными процессами в социальных сетях.

СЕКЦИЯ.
СУПЕРКОМПЬЮТЕРНЫЕ АРХИТЕКТУРЫ

«Эльбрус» сегодня: микропроцессоры, вычислительные комплексы и программное обеспечение

Ким А.К., Волконский В.Ю. (vol@mcst.ru), Груздов Ф.А., Сахин Ю.Х., Семенихин С.В., Фельдман В.М. — ОАО «ИНЭУМ им. И.С. Брука» (Москва)

Для сохранения экспоненциального роста производительности суперкомпьютерных систем и обеспечения возможности их эффективного использования, в ближайшие 10 лет придется решать новые сложные задачи. Американское агентство передовых оборонных научно-исследовательских разработок (DARPA) в программе «Повсеместные высокопроизводительные вычисления» (UHPC) определило следующие важнейшие задачи на период 2009–2018 годов: создание параллельной энергетически эффективной микропроцессорной архитектуры, обеспечение программируемости (снижение трудоемкости создания программ), существенный рост надежности и безопасности вычислительных систем. Создаваемые системы должны работать в широком диапазоне производительности — от терафлопсных встраиваемых систем до экзафлопсных суперкомпьютеров. Архитектурная линия российских микропроцессоров (МП) «Эльбрус», разработанная совместно с общим программным обеспечением (ОПО) «Эльбрус», позволяет решать поставленные в программе UHPC задачи.

В архитектуре МП линии «Эльбрус» используется явный параллелизм операций — распараллеливание программы выполняется оптимизирующим компилятором. За счет этого ядро МП «Эльбрус» может выполнять в несколько раз больше операций за один машинный такт по сравнению с другими современными архитектурами, не тратя энергию на распараллеливание при исполнении. В результате МП линии «Эльбрус» обладают большей логической скоростью (число операций, выполняемых за такт) и более высокой производительностью на единицу потребляемой энергии.

Универсальные МП линии «Эльбрус» позволяют использовать потенциал производительности с помощью оптимизирующих компиляторов, облегчая работу программистов за счет возможности использования языков высокого уровня. Программные средства динамической адаптации конкретной программы к аппаратным ресурсам обеспечивают более высокий коэффициент загрузки оборудования и, как следствие, почти в три раза большую логическую скорость на реальных программах по сравнению, например, с последними МП от Intel.

В архитектуре МП «Эльбрус» реализована аппаратно-программная защита программ и данных во время исполнения, что позволяет создать фундамент для построения безопасных систем широкого диапазона применений. На аппаратном уровне реализованы средства, обеспечивающие безопасное исполнение программ в едином виртуальном пространстве, исключающие возможность внедрения вредоносных кодов в программные системы и, за счет аппаратного контроля, позволяющие создавать надежные программы, выявляя наиболее сложные и неуловимые на других архитектурах ошибки.

В программе UHPC отмечается сложность перехода на новые архитектуры из-за проблем совместимости. В архитектурной линии «Эльбрус» заложены средства обеспечения эффективной и надежной аппаратно-программной совместимости с архитектурой x86 (x86-64). Оптимизация и накопление оптимизированных кодов, реализованные с помощью технологии скрытой динамической двоичной трансляции, обеспечивают более высокую производительность программ, представленных в кодах x86 (x86-64), при меньших затратах энергии.

Сегодня архитектура «Эльбрус» совместно с ОПО «Эльбрус» прошла проверку на трех поколениях микропроцессоров — МП «Эльбрус», система на кристалле (СнК) «Эльбрус-1С» и СнК «Эльбрус-2С+» — в составе вычислительных комплексов. МП «Эльбрус» [1] с производительностью 4,8 GFLOPS позволял создавать двухпроцессорные вычислительные комплексы на общей памяти и строить многомашинные системы на их основе, начиная с 2007 года. Созданная в 2010 году система «Эльбрус-1С» [2] производительностью 8 GFLOPS позволила

производить 4-процессорные одноплатные модули на общей памяти и вычислительные системы на их основе. Наконец, в 2011 году прошел успешные государственные испытания 6-ядерный гетерогенный МП «Эльбрус-2С+» (два универсальных ядра «Эльбрус» и 4 ядра DSP с архитектурой Мультикор) производительностью 28 GFLOPS. На базе этого МП выпускаются одноплатные многопроцессорные системы на общей памяти, на базе которых строятся мощные серверы с производительностью в несколько терафлопс, одноплатные двухпроцессорные встраиваемые системы, двухпроцессорные и однопроцессорные автоматизированные рабочие места, включая ноутбук и моноблок.

Вычислительные комплексы на базе МП с архитектурой «Эльбрус» оснащаются сертифицированным ОПО «Эльбрус», включающим операционную систему «Эльбрус», совместимую с ОС Linux, со средствами поддержки систем реального времени и средствами защиты от несанкционированного доступа. Средства разработки ПО обеспечивают эффективное распараллеливание программ, написанных на языках высокого уровня Си, Си++, Фортран, Java и др., на всех уровнях: параллелизм на уровне операций, векторный параллелизм, параллелизм потоков управления, параллелизм систем с распределенной памятью. ОПО «Эльбрус» включает средства поддержки пользовательского интерфейса, комплекс сервисных и пользовательских программ (СУБД, средства работы с гипертекстом, офисные пакеты, электронную почту и пр.), графические библиотеки и пакеты, высокопроизводительные математические и мультимедийные библиотеки. Эти средства поддерживают все возможности архитектуры «Эльбрус» и отвечают современным требованиям к программным системам индивидуального и коллективного пользования.

В 2012 году завершается разработка 4-ядерного универсального МП «Эльбрус-4С» с производительностью 64 GFLOPS, на базе которого могут создаваться 16-процессорные системы на общей памяти и мощные кластерные системы терафлопсного и петафлопсного диапазонов. Ведутся работы по созданию МП «Эльбрус-4С+» с производительностью 150 GFLOPS. В долгосрочных планах — создание МП «Эльбрус-8С» в 2017 году и «Эльбрус-16С» в 2019 году с производительностью 500 GFLOPS и 2 TFLOPS соответственно, что позволит до 2020 года создать российские вычислительные системы околоэкзафлопсного диапазона производительности на базе российских микропроцессоров.

Литература

1. Волконский Владимир, Груздов Федор, Ким Александр, Сахин Юлий. «Эльбрус» сегодня //Открытые системы. 2009. № 2.
2. Кузьминский Михаил. Куда идет «Эльбрус» //Открытые системы. 2011. № 7.

Актуальные проблемы создания и внедрения технологий суперкомпьютерного моделирования в науку и промышленность

Дерюгин Ю.Н., Костюков В.Е., Соловьев В.П., Шагалиев Р.М. (Sav@vniief.ru) — ФГУП «РФЯЦ-ВНИИЭФ» (Снежинск)

РФЯЦ-ВНИЭФ имеет более чем тридцатилетний опыт в области математического моделирования широкого спектра физических процессов на суперкомпьютерах. В РФЯЦ-ВНИИЭФ созданы математические методики и программные комплексы, предназначенные для комплексного моделирования сложных физических процессов на современных высокопроизводительных суперкомпьютерах с массовым параллелизмом. Ряд из них ориентированы на применение в интересах научных исследований и наукоемких расчетов высокотехнологичных гражданских отраслей промышленности.

В РФЯЦ-ВНИИЭФ реализуется комплексный подход к развитию суперкомпьютерных технологий, включающий в себя следующие основные направления: разработка усовершенствованных физико-математических моделей; конструирование численных методов, математических методик для решения многомерных нелинейных краевых задач;

создание эффективных методов распараллеливания задач на суперкомпьютерах с массовым параллелизмом; создание технологии комплексного сквозного моделирования (технология полномасштабных «компьютерных испытаний»); создание прикладных программных комплексов для суперкомпьютеров; разработка базового системного программного обеспечения для суперкомпьютеров и вычислительных центров коллективного пользования; проектирование и создание суперкомпьютеров различного класса.
Особое внимание уделяется работам по созданию отечественного программного обеспечения для имитационного моделирования с применением суперкомпьютеров и его внедрению в работы предприятий высокотехнологичных отраслей промышленности, проводимым в рамках реализации проекта Комиссии по модернизации и технологическому развитию экономики России — «Развитие суперкомпьютеров и грид-технологий». Излагаются научные подходы и достигнутые результаты, приводятся примеры использования созданного программного обеспечения для решения практических задач базовых отраслей промышленности (авиастроение, атомная энергетика, автомобилестроение, ракетно-космическая индустрия).
В докладе формулируются основные аспекты дальнейшего развития суперкомпьютерных технологий, ориентированных на применение перспективных архитектур. Представлены ключевые направления, основные задачи и проблемные вопросы в области развития технологии высокопроизводительных вычислений на базе суперкомпьютеров экзафлопсного класса.

Гибридный суперкомпьютер К-100: эволюция архитектур и эволюция пользователей

Дбар С.А. (ssa@kiam.ru), Жердева М.В. (zabrodin@kiam.ru), Лацис А.О. (lacis@kiam.ru), Орлов В.Л. (ovl@kiam.ru), Савельев Г.П. (sgp@kiam.ru), Смольянов Ю.П. (smol@kiam.ru), Храмцов М.Ю. (maximh@kiam.ru) — ФГБУН «ИПМ им. М.В. Келдыша РАН» (Москва)

Мировая суперкомпьютерная отрасль находится сегодня в фазе уверенного подъема: впечатляет динамика роста показателей в Top500; успешно преодолен петафлопсный рубеж; промышленность предлагает новые архитектурные и технические решения; пользователи предлагают новые задачи. Тем не менее есть все основания говорить о приближении серьезного кризиса, равных которому суперкомпьютерная отрасль в своей истории еще не знала. Кризис будет связан не с обострением проблем теплоотвода, энергетической эффективности вычислений или достижением физического предела скорости распространения сигнала в проводнике. Проблема, с которой предстоит справиться в ближайшее время, кроется в самой главной «детали» суперкомпьютера — в его пользователе. В чем же именно проблема? И проблема ли это или закономерный новый этап в развитии отрасли?
Все началось без малого четверть века назад, когда, благодаря прогрессу микроэлектроники, «транзисторный голод», терзавший разработчиков вычислительных систем, сменился «транзисторным изобилием» и быстро выяснились две важных вещи:

- архитектура «обычного» («фоннеймановского») процессора имеет предел эффективного использования «исходного материала» — транзисторов (бессмысленно объединять в одном «обычном» компьютере, скажем, миллиард транзисторов, как бессмысленно снабжать «обычный» автомобиль десятью тысячами колес);
- развитый до логического предела «обычный» процессор крайне неэффективно использует транзисторы, из которых он построен, разрыв между пиковым и реально достижимым быстродействием для реальных задач становится многократным.

Диалектика взаимодействия этих двух факторов заставила суперкомпьютерную отрасль пойти на смену архитектуры. Суперкомпьютеры вынуждены были стать системами массового параллелизма — место одного «суперпроцессора» заняли системы из многих тысяч отдельных компьютеров, связанных коммуникационной сетью. Суперкомпьютер впервые перестал быть похож на «обычный» компьютер с точки зрения общей логики его программирования.
Спустя еще 20 лет продолжающийся неуклонный прогресс в микроэлектронике привел

к необходимости более радикальной ревизии архитектуры. Впервые стало очевидным, что произошедший переход к многопроцессорным системам был не редким, однократным технологическим переворотом, а всего лишь первым шагом на пути все большего усложнения суперкомпьютерных архитектур, все большего их удаления от «обычного» процессора, в том числе — с точки зрения общей логики программирования. Сегодня мы наблюдаем второй шаг этого процесса — повсеместное распространение гибридных машин, таких как суперкомпьютер К-100, с ускорителями на графических процессорах. Когда случится и каким будет третий шаг? А четвертый?

Экстраполируя по двум (пока) точкам, можно сделать следующие выводы:

- Прикладное программирование традиционных машин массового выпуска, будучи, в громадном большинстве случаев, далеким от специфики высокопроизводительных вычислений, практически остается «обычным» (фоннеймановским). Именно это программирование изучают на первом курсе вузов будущие специалисты по высокопроизводительным вычислениям — прикладные программисты суперкомпьютеров.
- Доминирующие архитектуры суперкомпьютеров со временем усложняются, все более удаляясь от архитектуры «обычного» компьютера.
- Программирование суперкомпьютеров новой архитектуры усложняется пропорционально росту сложности архитектуры. За истекшую четверть века системные программисты не смогли надежно скрыть от прикладных программистов ни одного шага усложнения архитектуры суперкомпьютера.

Прогрессирующий разрыв между фоннеймановскими представлениями основной массы прикладных программистов о том, что такое компьютер, и объективно обусловленным, неизбежным и ускоряющимся усложнением суперкомпьютерных архитектур делает всю конструкцию современного суперкомпьютерного сообщества неустойчивой. Пройдет еще 3–4 года, будет сделан один, максимум — два шага в усложнении архитектуры суперкомпьютеров и вся «постройка» рухнет под гнетом сложности. Специалистов, связывающих свое профессиональное настоящее и будущее с суперкомпьютерной отраслью, в этой связи волнует вопрос о том, как именно это будет происходить. Можно ли этой разрушительной тенденции что-то противопоставить (вариант вопроса: не является ли грядущее разрушение отрасли в ее нынешнем виде неизбежным и, следовательно, вопрос о борьбе с ним бессмысленным)?

Ядро указанного противоречия — человеческий фактор. Понимая это, можно пытаться решать проблему с двух концов, а именно:

- автоматизировать прикладное программирование суперкомпьютера, позволив программисту использовать привычное и понятное фоннеймановское программирование;
- если это невозможно, то обучить прикладного программиста новому, гибридно-параллельному программированию.

Однако в действительности невозможно ни то, ни другое.

Продолжающиеся почти четверть века попытки автоматизировать прикладное программирование сначала параллельных, а затем гибридно-параллельных машин не привели к серьезным, прорывным результатам не по причине лени или глупости системных программистов, которые пытались (и все еще пытаются) решить эту задачу. Проблема имеет системный характер — для эффективного программирования принципиально новых архитектур изобразительные средства старых языков фоннеймановского типа в принципе не годятся. Требуется смена парадигмы, подобная той, которая потребовалась в 50-е годы при переходе от систем прямого аналогового моделирования физических процессов к их численному моделированию. При записи вычислительной процедуры фоннеймановским способом значительная часть информации о ней утрачивается и, вообще говоря, не может быть восстановлена никаким оптимизирующим компилятором — распараллеливателем, поскольку существует только как невысказанное знание в голове программиста.

Остается второй путь — обучить парадигме параллельного программирования непосред-

ственно прикладного программиста, заставить его высказать свое невысказанное знание на языке, понятном суперкомпьютеру. Первые 20 лет с того момента, когда суперкомпьютеры стали параллельными, отрасль шла именно по этому пути. Но с появлением гибридно-параллельных машин стало ясно, что сами парадигмы параллельного программирования, которым хотелось бы обучить прикладных программистов, не стабильны, будут отныне появляться регулярно, вместе с новыми архитектурами, неизбежно становясь при этом все более сложными и вычурными. Фоннеймановской парадигме почти 70 лет — за это время к ней привыкли все. Парадигме систем массового параллелизма на базе универсальных процессоров — почти 25. За это время к ней привыкли многие. Проживи она еще лет 40— необходимость автоматизации параллельного программирования отпала бы сама собой, но появление гибридно-параллельных машин с графическими процессорами не дает этих 40 лет. С новыми, более сложными архитектурами буквально завтра придет потребность в еще более новом параллельном программировании, которое сегодня и вообразить-то нельзя. Чтобы отслеживать этот все ускоряющийся и, несомненно, очень увлекательный процесс, прикладному программисту — пользователю суперкомпьютера пришлось бы постоянно тратить массу сил на изучение области знаний, далекой от его прямых профессиональных интересов. Опыт показывает, что очень многие пользователи на это не идут. Круг замкнулся.
Для дальнейшего развития в среднесрочной перспективе просматриваются два пути:

- системным программистам суперкомпьютерной отрасли все же удастся придумать принципиально новое программирование, которое позволит рано или поздно стабилизировать видимую прикладному программисту парадигму на десятилетия вперед;
- ближайшие годы станут началом процесса отмирания того вида пользователей суперкомпьютеров, к которому мы привыкли за последние 30–40 лет.

Второй путь наиболее вероятен, и речь идет, конечно же, о тех специалистах по численному моделированию в своих прикладных областях, которые сегодня самостоятельно проходят дорогу от уравнения в частных производных до разработки собственной прикладной программы для суперкомпьютера. Более или менее очевидно, что по крайней мере в России именно эта категория пользователей суперкомпьютеров составляет абсолютное большинство. Измученные переходом с «обычных» вычислений на параллельные, эти пользователи за 20 с лишним лет с трудом и не полностью смирились со своей участью. С перспективой снова и снова, каждые 4–5 лет, переживать этот кошмар большинство из них вряд ли согласится.
В результате сообщество пользователей опять, как когда-то в 50-е годы, разделится на «математиков — постановщиков задачи» и «программистов-кодировщиков». Вариантом такого разделения может стать расцвет производства готовых программных пакетов, изготавливаемых «кодировщиками» для «математиков» на регулярной, промышленной основе. Вопрос о том, сможет ли суперкомпьютерная отрасль найти внутри себя ресурсы на это весьма дорогостоящее производство, — тема отдельного исследования. В любом случае на краткосрочную перспективу можно прогнозировать значительный рост затрат на прикладное программирование суперкомпьютеров.

СЕКЦИЯ.
ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ СИСТЕМЫ ДЛЯ РЕШЕНИЯ ПРАКТИЧЕСКИХ ЗАДАЧ

Суперкомпьютерный комплекс МГУ: архитектура, пользователи, задачи

*Антонов А.С., Брызгалов П.А., Воеводин Вад.В., Воеводин Вл.В. (voevodin@parallel.ru), Жуматий С.А., Никитенко Д.А., Соболев С.И., Стефанов К.С. — НИВЦ МГУ им. М.В. Ломоносова (Москва)**

Суперкомпьютер «Ломоносов» был установлен в Московском государственном университете им. М.В. Ломоносова в 2009 году и по состоянию на осень 2012 года находился на 22-й позиции рейтинга Top500 с пиковой производительностью 1,7 PFLOPS и производительностью на тесте Linpack 901,9 TFLOPS. С момента запуска «Ломоносов» возглавляет список Top50 самых мощных компьютеров России.

Число вычислительных узлов x86	5104
Число вычислительных узлов GPU	1065
Число вычислительных узлов PowerXCell	30
Число процессоров x86	12 346
Число процессорных ядер x86	52 168
Число процессорных ядер GPU	954 240
Число типов вычислительных узлов	8
Основной тип вычислительных узлов	TB2-XN
Основные типы процессоров	Intel Xeon X5570 / X5670 NVIDIA X2070
Оперативная память	92 Тбайт
Системная сеть	QDR InfiniBand
Сервисная сеть	10Gigabit Ethernet
Управляющая сеть	Gigabit Ethernet
Специальная сеть	Сеть барьерной синхронизации и глобальных прерываний
Система хранения данных	Параллельная файловая система Lustre, файловая система NFS, система резервного копирования и архивирования данных
Операционная система	Clustrx T-Platforms Edition
Занимаемая площадь (вычислитель)	252 кв. м
Занимаемая площадь (всего)	1376 кв. м
Энергопотребление вычислителя	2,6 МВт
Общий вес оборудования	Более 75 т
Общая длина кабелей	Более 23,5 км
Объем/вес охлаждающей жидкости	50 т

В суперкомпьютере используются вычислительные узлы 8 видов и процессоры с различной архитектурой. В качестве классических многоядерных узлов x86 используются решения TB2-XN на базе четырехъядерных и шестиядерных процессоров Intel Xeon X5570 Nehalem и X5670 Westmere. Также используется платформа TB 1.1 с увеличенным объемом оперативной памяти и локальной дисковой памятью для выполнения специфических задач, требовательных к этим параметрам системы. Суперкомпьютерный комплекс содержит гибридные узлы TB2-TL на базе процессоров Intel Xeon и NVIDIA Tesla. Еще один тип узлов — платформы на базе многоядерного процессора PowerXCell 8i.

Все вычислительные узлы установки и систему хранения данных связывает высокоскоростная коммуникационная сеть QDR InfiniBand с пропускной способностью до 40 Гбит/с. В качестве дополнительных сетей используются 10Gigabit Ethernet и Gigabit Ethernet, а также выделенные сети поддержки коллективных коммуникаций.

Система бесперебойного энергоснабжения суперкомпьютера «Ломоносов» состоит из двух модульных источников бесперебойного питания Symmetra MW 1600 производства компании APC. Общая мощность системы бесперебойного энергоснабжения составляет 2800 кВт при уровне резервирования N+1. Охлаждение вычислительного комплекса строится с использованием внутрирядных кондиционеров. Шкафы с оборудованием размещаются в помещении машинного зала таким образом, чтобы образовались «горячие» и «холодные»

*Адаптированный вариант доклада авторов см. «Практика суперкомпьютера “Ломоносов”», «Открытые системы», № 07, 2012, www.osp.ru/os/2012/07/13017641/

коридоры. Кондиционеры забирают нагретый воздух из «горячего» коридора и подают охлажденный воздух в «холодный» коридор.

Для централизованного администрирования суперкомпьютера используется семейство специальных программных решений Clustrx, разработки холдинга «Т-Платформы». Решение включает в себя: ОС для серверов, построенную на базе Linux (CentOS 6.1); ОС вычислительных узлов, построенную на базе Linux; набор оптимизированных математических библиотек; комплект средств разработки; систему мониторинга и управления вычислительным комплексом Clustrx Watch; систему автоматического отключения оборудования; систему управления ресурсами на базе SLURM.

Для решения широкого спектра прикладных задач на суперкомпьютере установлен ряд

Процессоры узла	Число процессорных ядер на узел	Объем памяти узла (Гбайт)	Локальные диски	Число узлов
2 x Xeon 5570 2,93 Ггц	8	12	нет	4160
2 x Xeon 5570 2,93 Ггц	8	24	есть	260
2 x Xeon 5670 2,93 Ггц	12	24	нет	640
2 x Xeon 5670 2,93 Ггц	12	48	есть	40
2 x PowerXCell 8i 3,2 Ггц	18	16	нет	30
2 x Xeon E5630 2,53 Ггц, 2 x Tesla X2070	8+896 CUDA-ядер	12	нет	777
2 x Xeon E5630 2,53 Ггц, 2 x Tesla X2070	8+896 CUDA-ядер	24	нет	288
Xeon X7560 2,26 Ггц	128	2000	нет	1

программных пакетов: VASP, WIEN2k, Gaussian, CRYSTAL, MOLPRO, Turbomole, Accelrys Materials Studio, MesoProp, MOLCAS. Для разработки собственных приложений пользователи могут использовать компиляторы языков C/C++/Fortran с поддержкой стандарта OpenMP: GCC, Intel ICC/IFORT, PathScale, PGI. В состав математических библиотек системного ПО вычислительного комплекса входят ScaLAPACK, ATLAS, IMKL, AMCL, BLAS, LAPACK, FFTW, оптимизированные под архитектуру вычислительных узлов архитектуры x86; cuBLAS, cuFFT, MAGMA, cuSPARSE, CUSP, cuRAND, оптимизированные под архитектуру узлов GPU. В распоряжении пользователей имеются также средства отладки приложений: Intel VTune, Intel Trace Analyzer and Collector, Allinea DDT, RogueWave TotalView и ThreadSpotter.

Система хранения данных суперкомпьютера состоит из трех частей. Быстрое хранилище предназначено для проведения расчетов, построено на основе параллельной файловой системы lustre и доступно со всех узлов суперкомпьютера, включая вычислительные узлы, узлы доступа и узлы компиляции. Суммарный объем хранилища — до 500 Тбайт. Основное хранилище предназначено для хранения рабочих данных проектов пользователей. Это хранилище доступно по NFS с узлов доступа и компиляции. Общий объем этого хранилища— 312 Тбайт. Хранилище архивных данных расположено на ленте и имеет объем 580 Тбайт.

Возможности суперкомпьютерного комплекса МГУ используют более 600 научных групп из 24 подразделений МГУ, 35 институтов РАН, более 30 ведущих университетов России. Каждый день на «Ломоносове» в среднем выполняется около 700 сложных вычислительных задач— очередь пользовательских заданий держится на уровне 150–200 задач. Проведенный опрос показывает, что 51% научных групп для выполнения расчетов используют прикладные пакеты, 61% используют технологию MPI, 21% —OpenMP и 26% —комбинированный вариант MPI+OpenMP (часть научных групп использует в работе несколько подходов, поэтому сумма больше 100%). Для 71% пользователей принципиально важны арифметические вычисления с двойной точностью, для 18% — достаточно одинарной точности, у 11% пользователей вещественная арифметика не является доминирующей.

Совместной группой мехмата МГУ и ИПМ РАН получены уникальные результаты по численному моделированию формирования и развития концевых вихрей на сверхзвуковых режимах. С участием специалистов Научно-образовательного центра «Поисков, разведки и разработки

месторождений углеводородов» МГУ и российской компании «ГЕОЛАБ» решается ряд важных задач обработки сейсмических данных. Результаты в области криптографии получены совместно с группой с механико-математического факультета МГУ. Вычисления на суперкомпьютере «Ломоносов» были использованы сотрудниками НИВЦ МГУ в процессе разработки нового противоопухолевого лекарства на основе новых ингибиторов урокиназы. Изучается вопрос о связи динамики климата Земли с солнечной активностью. Получены первые результаты по воспроизведению внутригодовой изменчивости циркуляции вод Мирового океана с применением вихреразрешающей модели Мирового океана. В рамках совместных работ физического факультета МГУ, Института нефтехимического синтеза им. Топчиева РАН и Института исследований металлов общества Макса Планка (Дюссельдорф, Германия) исследуются методы создания новых полимеров. Физический факультет МГУ, Международный учебно-научный лазерный центр МГУ и Физический институт им. П.Н. Лебедева РАН ведут проект по созданию компактных лазерно-плазменных ускорителей заряженных частиц.

Суперкомпьютерный центр «Политехнический»: концепция и архитектура

Заборовский В.С. (vlad@neva.ru), Болдырев Ю.Я., Стрелец М.Ч. — СПбГПУ (Санкт-Петербург)

Суперкомпьютерные технологии (СКТ) становятся важным инструментом развития промышленного производства, существенно сокращая время и ресурсы, затрачиваемые на проектирование и испытания новых технических систем. Однако широкому внедрению СКТ в отечественную экономику мешают ряд факторов, наиболее значимыми из которых являются недостаточная техническая оснащенность, высокая стоимость владения аппаратным и программным обеспечением и отсутствие квалифицированных специалистов. Таким образом, организация в ведущих вузах страны эффективной системы подготовки кадров, обладающих знаниями и компетенциями в области применения СКТ, должна стать важной составляющей инновационной стратегии развития России.

В 2012 году Национальный исследовательский университет (НИУ) СПбГПУ разработал проект создания суперкомпьютерного центра (СКЦ «Политехнический»), ресурсы которого ориентированы на решение научно-технических задач. Особенностью таких задач является использование алгоритмов и численных методов, эффективная реализация которых требует применения современных технологий параллельных вычислений для компьютерных систем с различной архитектурой, включая гетерогенные и гибридные CPU/GPU, многоядерные микропроцессоры, многосокетные (архитектура ccNUMA с глобально адресуемой памятью) и реконфигурируемые вычислительные узлы (архитектура ПЛИС).

Проект создания СКЦ является частью плана инновационного развития СПбГПУ до 2020 года и предусматривает две фазы реализации: первая фаза — создание на площади СКЦ в новом здании научно-исследовательского корпуса вуза; вторая фаза — модернизация научно-образовательной структуры СПбГПУ с целью широкого внедрения технологий обучения и проведения научных исследований, ориентированных на использование СКТ для решения различных инженерно-технических задач на уровне, соответствующем международным стандартам качества инженерных разработок.

Базовые характеристики. Реализация первой фазы проекта включает в себя три этапа: формирование проектного облика суперкомпьютерного центра (2012 год), техническое проектирование (2013 год) и реализация проекта (2014 год). На стадии формирования проектного облика создаваемого суперкомпьютерного центра были решены следующие задачи: обоснована актуальность создания суперкомпьютерного центра для решения широкого класса научно-технических задач, носящих политехнический характер; сформулированы миссия и концепция СКЦ «Политехнический»; определены технические требования к программному и аппаратному обеспечению СКЦ, выполнение которых обеспечивает высокую производительность и масштабируемость численных решений для широкого спектра фундаментальных и прикладных, в том числе междисциплинарных, естественно-

научных задач (аэро- и гидроакустика, сопряженный тепломассоперенос, механика гомогенных и гетерогенных сред, проверка моделей, многофазное горение и т. п.), а также высокую эффективность функционирования систем обработки, передачи и хранения информации; выбраны базовые технические характеристики СКЦ по показателю вычислительной производительности (не менее 1 PFLOPS), использованию электроэнергии (PUE <=1,2), удельной энерговычислительной эффективности (более 1,5 GFLOPS/Вт); определены приоритеты прикладного использования ресурсов СКЦ в аспектах научной и образовательной деятельности.

На второй фазе реализации проекта (2015–2018 годы) планируется продолжить развитие СКЦ в направлении прикладного использования имеющихся вычислительных ресурсов, поддержки учебного процесса и организации научных исследований, в том числе: осуществлять разработку методов обучения, переподготовки кадров и проведение научных исследований, имеющих междисциплинарный характер на основе СКТ; проводить исследования по отработке технологий масштабирования ресурсов созданной суперкомпьютерной инфраструктуры и создания электронной компонентной базы с целью достижения удельных показателей эффективности вычислений, необходимых для перехода к экзафлопсному диапазону производительности; организовать вокруг СКЦ «кольцо» исследовательских центров, ориентированных на применение суперкомпьютеров и распространение опыта использования СКТ для решения междисциплинарных задач в различных областях науки и промышленности; создать защищенную среду обмена данными и управления доступом к вычислительным ресурсам СКЦ при взаимодействии пользователей с публичными информационными сервисами, доступными в сети Интернет; сформировать на базе СПбГПУ центр компетенции в области информационно-вычислительных и киберфизических систем, функционирующих с использованием сетецентрических технологий; обеспечить развитие научно-технического партнерства с учреждениями РАН, региональными, федеральными промышленными предприятиями и международными корпорациями в области применения СКТ в приоритетных областях науки и техники. Перечисленные работы планируется вести с использованием технологий облаков, сервисов дополненной реальности и визуализации.

Архитектура. Для достижения приведенных характеристик суперкомпьютерного центра и для отработки на его базе решений по переходу СКТ к экзафлопсному диапазону производительности предлагается использовать вычислительную среду, объединяющую:

- гибридные кластерные системы;
- компьютерные системы с глобально адресуемой памятью;
- реконфигурируемые вычислительные узлы на базе ПЛИС;
- систему хранения данных.

Комплекс программно-аппаратных средств будет находиться под контролем монитора загрузки, который интегрирован со средствами управления инженерными подсистемами СКЦ, обеспечивающими прецизионное охлаждение и бесперебойное питание вычислительных блоков, модулей хранения и сетевого обмена данных. Совместное управления потреблением энергоресурсов вычислительными и инженерными подсистемами позволит обеспечить для оборудования СКЦ наилучший показатель PUE. В свою очередь, снижение совокупной стоимости владения ресурсами позволит повысить экономическую эффективность использования СКТ для решения широкого класса вычислительных задач в режиме аутсорсинга и организации доступа к информационным сервисам на принципах SaaS и IaaS. Использование современных технологий виртуализации в рамках облаков для динамической конфигурации ресурсов обеспечит работу СКЦ в соответствии со стандартами экологичности, безопасности и энерговычислительной эффективности, что позволит аттестовать созданный комплекс технических средств по критериям Green500 (www.green500.org).

СКЦ планируется разместить в строящемся научно-исследовательском корпусе, инженерная инфраструктура которого обеспечит выделение мощности для системы электропитания до 2 МВт, средства удаленного сетевого мониторинга ресурсов, комплексную систему безопасности, защиты информации, физического контроля доступа и тушения пожаров. Основным вычислительным комплексом, входящим в инфраструктуру СКЦ, будет кластерная

вычислительная система, состоящая из гибридных узлов, ориентированных на решение задач, характерных для научных и инженерных расчетов с использованием коммерческих и свободно распространяемых пакетов прикладных программ. Вычислительный кластер будет состоять из более чем 2 тыс. вычислительных узлов, каждый из которых может содержать два/четыре многоядерных процессора, а часть дополнительно оборудована ускорителями на базе графических процессоров. Суммарное количество процессорных ядер будет более 40 тыс., а ядер графических ускорителей — более 160 тыс. Расчетная пиковая производительность кластерной системы без учета графических ускорителей составит более 420 TFLOPS, а с учетом графических ускорителей — примерно 540 TFLOPS. Другим звеном будет вычислительная система, состоящая из узлов, имеющих глобально адресуемую память, что даст возможность эффективно решать задачи, требующие большого объема оперативной памяти и программных средств без поддержки протокола MPI. В составе СКЦ будет функционировать вычислительная подсистема на базе реконфигурируемых модулей ПЛИС. Эта подсистема предназначена для реализации алгоритмов параллельных расчетов, требующих большого объема вычислений сложных математических функций (экспонент, нахождения целочисленных корней алгебраических полиномов и пр.).
Для обеспечения работы всех вычислительных подсистем и обработки данных моделирования в СКЦ будет установлена высокопроизводительная система визуализации расчетов и виртуального прототипирования созданных моделей. Общий объем системы хранения составит 700 Тбайт.

***

Реализация проекта СКЦ «Политехнический» обеспечит решение следующих задач:

- создание нового облика научно-образовательной среды вуза XXI века, который будет формироваться на основе отечественного опыта развития политехнической школы фундаментальных инженерных знаний и СКТ мирового уровня;
- создание в Санкт-Петербурге и Северо-Западном регионе России суперкомпьютерной инфраструктуры нового технологического уклада, обеспечивающей ускоренное внедрение новейших достижений фундаментальной науки в промышленность, образование, здравоохранение и социальную сферу;
- формирование научно-технических заделов по развитию отечественных СКТ и их применению в интересах наукоемких секторов национальной экономики, системы подготовки и переподготовки кадров, которые обладают знаниями и компетенциями, необходимыми для внедрения инноваций в приоритетные сферы научных исследований и промышленного производства.

Некоторые вопросы использования высокопроизводительных кластеров для решения задач корабельной гидродинамики

Лобачев М.П. (lobachevm@mail.ru), Овчинников Н.А., Пустошный А.В. — ФГУП «ЦНИИ им. А.Н. Крылова» (Санкт-Петербург)

Суперкомпьютеры позиционируются как системы для решения задач большой и сверхбольшой размерности, а каковы сегодня реальные потребности в решении таких задач в судостроении? Какие проблемы возникают на пути использования суперкомпьютеров в судостроении?
Решение большинства гидродинамических задач в судостроении сегодня осуществляется на основе использования уравнений Рейнольдса с моделированием турбулентности полуэмпирическими моделями типа k-ε или k-ω SST. Для некоторых задач используются гибридные методы моделирования турбулентности типа DES. Характерная размерность задач— от одного до нескольких сотен миллионов расчетных ячеек, причем значительная часть задач нестационарны.
Эллиптический характер уравнений требует использования данных (значения переменных

и значения коэффициентов уравнений) во всем расчетном объеме. Применение маршевых методов, при которых последовательно используется информация только на нескольких сечениях, здесь невозможно. Это приводит к постоянной работе с массивами данных большой размерности, причем обращение к памяти имеет случайный и нерегулярный характер. Именно это обстоятельство приводит к тому, что реальная производительность используемых суперкомпьютеров оказывается в пределах 12–15% от пиковой, а не порядка 80%, как на тесте Linpack. В результате при проведении параллельных вычислений на таких задачах достаточно быстро наступает ситуация, когда увеличение числа используемых ядер не приводит к росту скорости счета. Так, на протестированных процессорах и конфигурациях признано нецелесообразным выделение менее 50 тыс. расчетных ячеек на одно ядро процессора. Эта величина также зависит от используемого программного обеспечения, в частности от реализации библиотеки MPI. Есть некоторая зависимость этой величины и от размерности задачи, однако в целом эти колебания укладываются в диапазон 30–100 тыс. расчетных ячеек на ядро. Увеличение времени счета может быть вызвано задержками, связанными не только с ожиданием передачи данных между узлами, но и с ожиданием данных из оперативной памяти внутри узла. Поэтому преимуществом обладают системы, обеспечивающие большую пропускную способность памяти.

В таблице приведены данные по размерности расчетных сеток, времени расчета одного шага (для стационарных задач — времени расчета одной итерации), времени решения задачи в целом (до сходимости) для нескольких наиболее типичных задач корабельной гидродинамики. Время решения задачи в целом также зависит от того, требуется ли только получение усредненных характеристик или же требуется анализ нестационарных переменных (сил, моментов, скоростей и т. п.). В последнем случае требуются дополнительные расчеты для накопления необходимой для выполнения анализа длины выборки.

Задача	**Размерность сетки, (млн ячеек)**	**Исполь-зуемое число ядер**	**Время расчета одной итерации (с.)**	**Время решения задачи, (ч.)**	**Требуемое время решения, (ч.)**
Расчет обтекания корпуса	8,2	144	2	1,2	3
Расчет обтекания гребного винта в однородном потоке	15,8	336	28,5	23,75	0,5
Расчет обтекания гребного винта за корпусом судна	17	336	75	105	8
Расчет многолопастного, многорядного движителя за корпусом судна	131	2328	107,5	150	240
Расчет течения в трубе с локальным изменением сечения (определение пульсаций сил в заданном диапазоне частот)	12,4	264	36,5	234	300

Таблица. Данные по расчетам

Расчеты выполнялись на кластере пиковой производительностью 21,12 TFLOPS (17,38 — на тесте Linpack) из 100 двухпроцессорных узлов с 12-ядерными процессорами, объединенных сетью QDR Infiniband. Количество используемых ядер определялось ограничением 50 тыс. ячеек на ядро с округлением до числа, кратного 24 ядрам, чтобы полностью занимать расчетный узел. Фактически в таблице указано минимальное время решения задачи, а под требуемым временем решения понимается максимальное время расчета, при котором данные

расчеты могут быть использованы в текущей практике проектирования. Задачи из первой и третьей строки таблицы относятся к разряду постоянно проводимых работ, для обеспечения которых существуют альтернативные методы (узкоспециализированные расчетные или экспериментальные), однако давно назрела необходимость перехода на более точные методы. Задача из строки 4 относится к уникальным, для которых время проведения проектных работ пока еще может быть значительным, но в дальнейшем все равно потребуется ускорение счета. Задача из строки 5 возникла при исследовательской работе, выполнявшейся в обеспечение перспективных проектных проработок, причем при использовании альтернативных методов решения возникли значительные трудности, поэтому столь большое время расчета пока оказалось приемлемым. Таким образом, можно констатировать, что требуется увеличение скорости счета, причем для ряда задач — весьма значительное, но достигаемое не за счет увеличения количества используемых ядер. Следует также отметить, что применение для ускорения счета GPU на таких задачах невозможно из-за использования исключительно неявных алгоритмов, на которых GPU неэффективны.

Из таблицы хорошо видно, что современные методы расчета с использованием суперкомпьютеров традиционной архитектуры пока не могут быть полностью внедрены в текущую практику проектирования. Их использование в основном ограничено уникальными разработками, на выполнение которых отводится больше времени. Следует отметить, что при решении задач из строк 2 и 3 таблицы они все же используются, но как дополнительные для решения задач, которые не могут быть решены традиционными методами (узкоспециализированными расчетными или экспериментальными). Такая необходимость сегодня существует, и, несмотря на невозможность в этом случае использования современных численных методов при проведении основного массива проектных работ, они уже повсеместно внедряются для решения отдельных специальных вопросов. В ЦНИИ им. акад. А.Н. Крылова уже накоплен достаточно большой опыт такого внедрения.

При решении задач малой размерности (до 10 млн ячеек, когда расчет производится на менее чем 10 узлах) для связи между расчетными узлами возможно использование сети Gigabit Ethernet, но для задач большей размерности необходимо использовать более скоростные сети, например QDR InfiniBand. При увеличении количества ячеек, приходящихся на одно ядро, время решения задачи увеличивается, однако при этом также уменьшается количество коммуникаций между параллельными процессами и, как следствие, меньше времени тратится на пересылку данных. При решении нескольких задач, каждая из которых имеет расчетную сетку в пределах 10–50 млн, оптимальным оказывается выделение количества ядер, приходящихся на одну задачу, из условия 1 млн ячеек на ядро. При этом время, затрачиваемое на решение одной задачи, оказывается больше, но суммарное время решения всех задач сокращается. Задачи, полностью занимающие один расчетный узел, дополнительно получают преимущества от использования более быстрого механизма передачи сообщений через общую память, без использования сетевого соединения. Правильное распределение нагрузки при решении нескольких задач на кластере позволяет повысить эффективность использования ресурсов.

Для генерации расчетной сетки с использованием большинства существующих коммерческих генераторов сеток требуется не кластер, а SMP — компьютер с общей памятью. Предельная размерность расчетной сетки, которая может быть построена на таком компьютере, определяется общим объемом ее памяти. Для построения расчетной сетки в среднем требуется 1 Гбайт на 1 млн расчетных ячеек. В случае нехватки оперативной памяти будет производиться запись на жесткий диск, что существенно снижает скорость построения расчетной сетки. Для наших задач мы используем четыре машины, имеющие по 256 Гбайт оперативной памяти.

***

В целом можно констатировать, что для решения текущих задач корабельной гидродинамики при сохранении существующей архитектуры и скорости сетей, используемых для обмена между расчетными узлами, петафлопсные компьютеры традиционной архитектуры не требуются — все равно возникают ограничения, возникающие при увеличении количества используемых

ядер. Требуется переход к суперкомпьютерам другой архитектуры. Исходя из практического опыта решения реальных задач можно констатировать, что нужны небольшие дешевые системы с производительностью 1–5 TFLOPS, не требующие дополнительной инфраструктуры. Кроме этого, нужны суперкомпьютеры с быстродействием 20–250 TFLOPS, необходимые для решения больших задач и отладки методик и технологий численного эксперимента.

Суперкомпьютерное моделирования задач многофазной фильтрации

Афанасьев А.А. (afanasyev@imec.msu.ru) — ЗАО «Т-Сервисы» (Москва)

Сегодня для уменьшения выбросов парниковых газов в атмосферу интенсивно исследуется возможность надежного захоронения углекислого газа в проницаемых недрах Земли, в частности в водонасыщенных пластах. Так как углекислый газ легче воды, то из-за эффекта плавучести он может всплыть к поверхности и вернуться в атмосферу. Для надежного прогнозирования последствий захоронения углекислого газа необходимы сложные трехмерные модели, учитывающие реальные геологические параметры гетерогенного проницаемого резервуара и многофазную неизотермическую специфику течений воды и углекислого газа. Как правило, подобные многопараметрические модели могут использоваться только совместно с прямыми численными расчетами на суперкомпьютерах.
В работе представлены результаты численного моделирования захоронения углекислого газа в формации Johansen, расположенной в Северном море. В параллельных расчетах используется реальная трехмерная геологическая модель проницаемого коллектора, учитывающая гетерогенные свойства пород, значительное изменение глубины залегания пластов и геологические разломы. Проведена серия расчетов при различных интенсивностях нагнетания углекислого газа и различных расположениях скважины, через которую происходит закачка. Определено влияние различных механических процессов на количество углекислого газа, удерживающегося в пластах. Показано, что за счет более удачного расположения скважины можно значительно повысить эффективность захоронения.

Суперкомпьютеры для решения задач газовой динамики в узлах перспективных авиадвигателей

Бендерский Л.А., Виноградов В.А, Жемуранова Л.Д., Любимов Д.А., Ляшенко В.П., Макаров А.Ю., Потехина И.В., Степанов В.А. (step@ciam.ru), Строкин В.Н., Шихман Ю.М. — ФГУП «Центральный институт авиационного моторостроения им. П.И. Баранова» (Москва)

В докладе продемонстрировано решение научно-практических задач численного моделирования аэрогазодинамических и теплофизических процессов для проектирования различных узлов авиационных двигателей: воздухозаборников, переходных каналов, камер сгорания. Моделирование проведено с использованием коммерческих пакетов ESI Group (Fastran и ACE), Ansys и авторского кода Jet3D. Для расчета таких задач использовались четырехсокетные 48-ядерные узлы на серверной плате Supermicro с четырьмя 12-ядерными процессорами Opteron 6174 (AMD MagnyCours/2,2 ГГц), с общей памятью 128 Гбайт, памятью на дисках 1 Тбайт и сетевыми адаптерами. Соединение между сокетами — 8-битовые линки HyperTransport с пропускной способностью 6,4 Гбайт/с, пропускная способность одного чип-сокета с модулем памяти — 28,8 Гбайт/с, односторонняя пропускная способность интерфейса PCI-экспресс 2.0 (16 бит) — 8 Гбайт/с, односторонняя пропускная способность линка сети Infiniband QDR — 4 Гбайт/с. Чип-сокеты внутри одного сокета соединены 24-битовым интерфейсом HyperTransport — 19,2 Гбайт/с. Пиковая производительность одного ядра на скалярных операциях — 8,8 GFLOPS, а всего 12-ядерного микропроцессора — 105,6 GFLOPS.
На предприятии проведено численное моделирование обтекания пространственного высокоэффективного входного устройства в компоновке с фюзеляжем четырехдвигательного

сверхзвукового пассажирского самолета (СПС) с крейсерским режимом полета M = 1,8, включающего два двухканальных сверхзвуковых воздухозаборника (ВЗ) с криволинейными дозвуковыми диффузорами большой кривизны. Получены поля чисел Маха в поперечных сечениях ВЗ и дроссельные характеристики ВЗ (зависимости коэффициента полного давления от коэффициента входа) для каждого канала ВЗ. Расчет проводился интегрированием системы уравнений Навье – Стокса с помощью пакета ESI Group FASTRAN. Использовалась численная схема второго порядка аппроксимации с использованием схемы Roe. Для описания турбулентных характеристик в разных областях течения использовалась двухпараметрическая «K-ε»-модель турбулентности. Как показали исследования, подобные численные схемы эффективны при расчете высокоградиентных сверхзвуковых течений при наличии скачков уплотнений.

Решение задачи проводилось при использовании многоблочной регулярной сетки из прямоугольных ячеек общим количеством до $\approx 19 \times 10^6$ ячеек, адаптированной для расчета вязких пристеночных слоев для лучшего разрешения пограничного слоя и с достаточной густотой в ядре потока. Расчетная сетка строилась с использованием программы ESI Group GEOM, входящей в состав всего пакета ESI Group. Расчеты проводились на 2 узлах. Расчет одной точки дроссельной характеристики занимал около 45 часов. Задача была распределена с помощью утилиты MDICE.

Также были проведены подробные исследования дроссельных характеристик изолированных одноканальных пространственных ВЗ. Получены картины чисел Маха и коэффициента полного давления в продольных и выходных сечениях ВЗ. Расчеты проводились с использованием решателя FASTRAN, но расчетная сетка была уменьшена до 5 млн ячеек. Расчеты позволили выбрать оптимальную конфигурацию ВЗ и получить высокую эффективность сжатия и низкие уровни дополнительного аэродинамического сопротивления.

Результаты моделирования процессов распространения и смешения струй в канале модульного гиперзвукового воздухозаборника были получены при вдуве газа (метана) в следе за пилоном, установленным перед ВЗ. Продемонстрирована высокая эффективность смешения при устойчивости течения в канале ВЗ. Показаны поля концентраций метана в поперечных и продольных сечениях ВЗ. Расчеты подтверждены экспериментальными исследованиями.

Исследовано входное устройство высокоскоростного воздушно-реактивного двигателя в компоновке с корпусом ЛА при числе Маха M = 2,5–6. Получены средние значения чисел Маха и коэффициента полного давления в выходных сечениях. Проведено сравнение полученных результатов с данными эксперимента и расчетами с помощью пакета FLUENT. Показано, что пакет FASTRAN демонстрирует лучшее согласование с экспериментом.

С помощью решателя ACE в рамках пакета ESI GROUP проводились исследования стационарных и нестационарных эффектов при обтекании переходных каналов в авиационных двигателях. Показано влияние вдува синтетических струй в модельный диффузорный канал. Вдув позволил уменьшить потери в диффузоре. Эксперимент показал аналогичный эффект — уменьшение потерь полного давления при различных мощностях излучателя. Проведено сравнение с экспериментальной картиной, полученной с помощью метода PIV.

Получены результаты по моделированию влияния пластинчатых завихрителей на течение в цилиндрических диффузорных переходных каналах перспективных авиадвигателей. Применение оптимальных по форме наборов пластинчатых завихрителей позволило снизить потери в диффузоре почти вдвое, что согласуется с данными экспериментальных исследований модельного диффузора.

Расчеты процессов в отсеке полноразмерной укороченной камеры сгорания ТРД проведены с использованием решателя ACE. Получение экспериментальных данных затруднено из-за высоких энергетических и финансовых затрат. Показаны результаты по моделированию процесса горения углеводородного топлива в 30-градусном отсеке камеры сгорания с использованием 4-реакционной схемы. В расчете выявлена высокая полнота сгорания на малых длинах и низкие выбросы вредных примесей NOx, CO и др. Показаны поля температур в поперечном и продольном сечении камеры сгорания. Приведены кривые испарения и

выгорания топлива. Проведены также расчеты с помощью пакета FLUENT, которые подтвердили данные, полученные с помощью решателя ACE с использованием расчетной сетки до 100 млн ячеек. Таким образом, были решены сложные прикладные задачи по моделированию рабочих процессов в узлах авиационных двигателей с использованием суперкомпьютеров.

В докладе представлены результаты расчетов, выполненных с помощью комбинированного RANS/ILES-метода высокого разрешения (9-й порядок), реализованного в авторском коде Jet3D. Код был написан на языке Фортран, а расчеты выполнялись на структурированных многоблочных сетках. Распараллеливание осуществлялось с помощью Open MP. Это оправданно при использовании разностных схем высоких порядков с протяженным конечно-разностным шаблоном (7 или 11 точек). При использовании MPI много времени тратится на обмен. Каждый расчет выполнялся на одном узле.

Проведено исследование отрывного турбулентного течения в межкомпрессорном диффузоре с диффузорностью F = 2,07 (ReH = 3,8–5,5×10^5). Расчетная сетка — 0,8×10^6 ячеек. Время счета — 80 часов. Было установлено, что двумерное течение на входе преобразуется в трехмерное в выходном сечении диффузора. Это является следствием локального характера отрыва в азимутальном направлении. При этом количество локальных отрывных зон зависит от времени осреднения и перепада давления в канале. Изучено влияние неоднородности полного давления на входе в канал на течение в межкомпрессорном кольцевом диффузоре.

Исследовано отрывное турбулентное течение в межтурбинном диффузоре с диффузорностью F = 2,7 (ReH = 1,1–21,0×10^5). Расчетная сетка — 0,8×10^6 ячеек. Время счета — 120 ч. В этом случае течение в выходном сечении диффузора было полностью турбулентным и с развитым отрывом.

Таким образом, эффективный RANS/ILES-метод высокого разрешения был применен для расчета сложных турбулентных течений в элементах ТРД. Использование распараллеленного кода и кластеров позволило выполнять до нескольких сотен различных вариантов массовых параметрических исследований. Были исследованы: турбулентные течения в сложных диффузорах (в том числе и активное управление течениями в них), нестационарные явления в сверхзвуковом воздухозаборнике при различных режимах работы двигателя, нестационарные характеристики течения на входе в двигатель, турбулентные струи из сопел различных типов и на различных режимах (в том числе и сверхзвуковые со скачками уплотнения), влияние компоновки на течение в струе. Это задачу очень сложно исследовать экспериментально без применения суперкомпьютеров.

Во всех представленных задачах точность превышала достигнутую с помощью RANS-методов (часть задач вообще невозможно решить с помощью RANS). Результаты расчетов хорошо совпадают с известными экспериментальными данными и дополняют их.

СЕКЦИЯ.
КАДРЫ ДЛЯ ИНДУСТРИИ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Международные образовательные стандарты — методологический базис системы ИТ-образования

Сухомлин В.А. (sukhomlin@mail.ru) — МГУ им. М.В. Ломоносова (Москва)

В условиях глобализации экономики большое значение для подготовки кадров имеет выработка международных рекомендаций, обладающих высоким уровнем консенсуса в профессиональной среде и служащих ориентиром для университетов и вузов. Такого рода рекомендации должны систематизировать и унифицировать требования практики к выпускникам вузов и к соответствующим образовательным программам, учитывать достижения и тенденции развития предметной области, обобщать лучшую образовательную практику, служить эффективным инструментом построения актуальных образовательных программ, единого образовательного пространства.

Ответственность за решение задачи формирования таких ориентиров-рекомендаций в виде типовых учебных программ (куррикулумов, curriculum) взяли на себя ведущие международные профессиональные организации: ACM (Association for Computing Machinery) и IEEE-CS (Computer Society of the IEEE), ведущие такую работу начиная с 60-х годов.

В 1998 году вновь созданная объединенная группа специалистов под эгидой ACM и IEEE-CS приступила к разработке обновленной версии CC — Computing Curricula 2001 (CC 2001) [1]. Разработчикам этого документа уже на стадии анализа стало ясно, что за последнее десятилетие область ИТ претерпела столь значительные изменения, что для ее адекватного представления в академическом пространстве необходимо разработать целую систему куррикулумов. С этого времени процесс разработки стандартов куррикулумов приобрел непрерывный характер, и к середине первого десятилетия этого века был разработан целостный набор куррикулумов, венцом которого стал документ Computing Curricula 2005 (CC2005) [2], имеющий общеметодологическое назначение.

За последнее пятилетие практически все описанные в CC2005 куррикулумы были переработаны и вышли в новых редакциях. Архитектура современного стека куррикулумов показана на рисунке. Подробное описание основных документов этого стека приведено в работе [3].

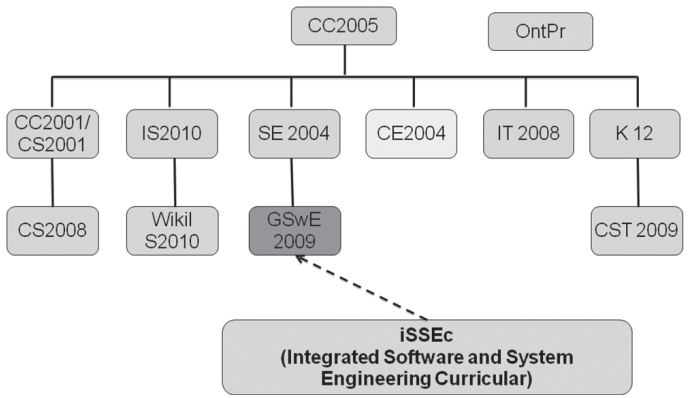

Рисунок. Современная система международных стандартов в области ИТ

Анализ методологической базы построения учебных программ по направлению ИТ позволяет увидеть общую картину процесса стандартизации учебных программ, выявить основные тенденции, закономерности, характерные особенности этого процесса. В частности, к

общим принципам его реализации можно отнести такие стороны, как целенаправленность, системность, концептуальную целостность, модульность, консорциумностный подход при создании куррикулумов, оценку их качества на основе консенсуса. Характерные черты стандартов куррикулумов:

- единая структура построения и единый понятийный контекст;
- знание-ориентированность — спецификация структуры и собственно объемов знаний (body of knowledge) по профилям подготовки (до уровня тем/подтем) является основным содержанием любого куррикулума;
- единый способ структурирования и представления объемов знаний в виде трех-четырехуровневой иерархической структуры: на верхнем уровне иерархии располагаются предметные области (areas) — самые крупные части объема знаний, которые подразделяются на разделы или модули знаний (units), последние, в свою очередь, разбиваются на темы (topics), в некоторых случаях темы делятся на подтемы (subtopics);
- концепция ядра — выделение в объемах знаний минимально необходимого образовательного содержания, реализация которого во всех учебных программах обеспечивает единство образовательного пространства, мобильность учащихся в рамках профиля или всего направления, гарантию качественности базовой подготовки;
- четкая спецификация профессиональных характеристик профилей, системы целей обучения, итоговых профессиональных характеристик выпускников;
- рекомендации методического характера по диверсификации направлений подготовки, составлению учебных планов, компоновки курсов из модулей знаний в соответствии с выбранной педагогической стратегией реализации учебной программы, организации профессиональной практики, реализации процессов обучения;
- описание учебных курсов и пакетов курсов для различных педагогических стратегий реализации куррикулумов.

Опыт создания и последующего развития образовательного направления 010300 «Фундаментальная информатика и информационные технологии» (первоначальное название направления — «Информационные технологии») убедительно подтвердил эффективность использования международных стандартов в разработке основных образовательных программ в области ИТ, являющихся общепринятой методологической базой современной системы ИТ-образования.

Литература

1. Computing Curricula 2001. Computer Science Volume. Association for Computing Machinery and Computer Society of IEEE. http://www.acm.org/education/cc2001/final

2. Computing Curricula 2005 (CC2005). Association for Computing Machinery and Computer Society of IEEE.

3. Сухомлин В.А. Международные образовательные стандарты в области информационных технологий // Прикладная информатика. 2012. № 1(37), январь-февраль. С. 33–54.

Образовательный проект «Интернет-университет суперкомпьютерных технологий»

Баркалов К.А. (KonstantinBarkalov@yandex.ru), Гергель В.П. (gergel@unn.ru) — ННГУ им. Н.И. Лобачевского (Нижний Новгород); Воеводин Вл.В. (voevodin@parallel.ru) — Научно-исследовательский вычислительный центр МГУ (Москва); Шкред А.В. (anatoli@shkred.ru) — Национальный открытый университет «ИНТУИТ» (Москва)

Появление вычислительных кластеров и многоядерных компьютеров позволило значительно расширить масштаб применения высокопроизводительных вычислений. Многие инженерные и технологические процессы в машиностроении, фармацевтике, энергетике, биоинженерии и нанотехнологиях теперь опираются на использование быстродействующих вычислительных

систем. То, что раньше можно было решить только на дорогих, а потому для многих практически недоступных суперкомпьютерах, сегодня можно сделать с помощью недорогих кластерных систем. Широкое использование параллельных вычислительных технологий стало характерной чертой нашего времени.

Сегодня на первый план выходит задача разработки нового и адаптации существующего научно-методического обеспечения для массовой подготовки специалистов в области суперкомпьютерных технологий и параллельного программирования, в особенности — дистанционно, с использованием возможностей Интернета.

Целью образовательного проекта «Интернет-университет суперкомпьютерных технологий» (www.hpcu.ru) является организация массовой подготовки специалистов в области высокопроизводительных вычислений и суперкомпьютерных технологий с активным использованием возможностей Интернета. Для достижения поставленных целей осуществляется решение следующих задач:

- разработка научно-методического обеспечения для проведения занятий в области суперкомпьютерных технологий;
- подготовка новых и адаптация существующих курсов лекций;
- проведение обучения слушателей при проведении занятий в форме интернет-видеоконференций.

Партнеры проекта: «ИНТУИТ.РУ»; издательство «Открытые системы»; НИВЦ МГУ; Нижегородский государственный университет им. Н.И. Лобачевского (ННГУ).

Образовательная деятельность университета ориентирована на обучение широкого круга обучаемых (студенты, специалисты, преподаватели) и предусматривает наличие различных направлений подготовки для учета разных профессиональных требований в области суперкомпьютерных технологий (пользователи, программисты, инженеры).

Согласование массовости обучения и качества получаемого образования обеспечивается за счет привлечения к деятельности университета ведущих специалистов страны в области суперкомпьютерных технологий, активного использования современных ИТ-технологий для организации учебного процесса, применения двухуровневой системы обучения:

- базовый уровень подготовки ориентирован на самый широкий круг обучаемых (студенты, специалисты, преподаватели) и организуется в форме дополнительного образования на основе технологий дистанционного обучения с использованием Интернета (включая проведение учебных занятий в виде видеоконференций), планируемая длительность данного вида обучения — 1 год;
- профильный (углубленный) уровень подготовки организуется в опорных образовательных центрах университетов — участников проекта, длительность такого обучения регулируется образовательными программами, реализуемыми в соответствующих образовательных центрах.

Слушателям, успешно прошедшим базовый уровень подготовки, выдается сертификат университета. Слушатели, прошедшие обучение в образовательных центрах университета, получают государственные дипломы университетов-участников, в которых организованы соответствующие образовательные центры.

Процесс обучения в рамках университета формируется на основе технологий дистанционного образования (представление учебных материалов в Интернете, модульное представление учебного материала, автоматизированное тестирование), а для организации учебного процесса активно используется методика обучения, применяемая в Интернет-университете информационных технологий (www.intuit.ru). Для представления учебных материалов наряду с традиционным гипертекстовым форматом широко используются видеоматериалы.

Основной составляющей обучения является проведение учебных занятий в форме видеоконференций, позволяющих обеспечить на новой технологической основе возможности классического очного обучения (изложение учебного материала преподавателем, опрос обучаемых, организацию самостоятельной работы под управлением преподавателя). Технологической основой для организации видеозанятий является коммуникационный

сервер Microsoft Office Communication Server, а в качестве клиента служит программа Microsoft Office Live Meeting, позволяющая преподавателю: организовать двустороннюю аудио- и видеосвязь с аудиторией; получать вопросы от слушателей как в текстовом, так и в звуковом виде; показать слушателям любую программу, выполняемую на компьютере преподавателя (например, презентацию лекции); показать слушателям «рабочий стол» компьютера; дать задание для самостоятельной работы; организовать запись лекции для последующего просмотра и использования.

Самостоятельная работа слушателей университета и возможность проведения вычислительных экспериментов в процессе обучения поддерживаются предоставлением доступа к суперкомпьютерным центрам университетов — исполнителей проекта. В рамках выполнения проекта предусматривается возможность регулярного проведения очных семинаров-школ (например, в рамках научно-технических конференций суперкомпьютерной тематики).

***

В результате выполнения проекта будет разработано научно-методическое обеспечение системы дистанционного образования в области суперкомпьютерных технологий с активным использованием глобальной сети Интернет. Разработанное научно-методическое обеспечение апробировано при проведении дистанционного обучения широкого круга слушателей (студенты, специалисты, преподаватели). Количество слушателей, зарегистрировавшихся на сайте проекта, — около 2 тыс. человек.

СОДЕРЖАНИЕ

Тезисы докладов Третьего Московского суперкомпьютерного
форума, Москва, 2012, 48 с.,
Формат 60x90 1/16. Печ. л. 3,25
Тираж 400. Заказ № 1026 от 26.11.2012

Отпечатано в типографии
ЗАО "Новые печатные технологии"
тел.: + 7 (495) 223-92-00
info@web2book.ru, www.web2book.ru