

# ПРОШЛОЕ И НАСТОЯЩЕЕ РОССИЙСКИХ СУПЕРКОМПЬЮТЕРОВ

## ПО МАТЕРИАЛАМ ДОКЛАДОВ ТРЕТЬЕГО НАЦИОНАЛЬНОГО СУПЕРКОМПЬЮТЕРНОГО ФОРУМА

М. Шейкин max.shaking@ya.ru

В конце ноября 2014 года Институт программных систем РАН им. А.К.Айламазяна (г. Переславль-Залесский) принимал гостей третьего Национального суперкомпьютерного форума – самого важного и глобального мероприятия для российской суперкомпьютерной отрасли. Прозвучавшие на конференции доклады были посвящены как глобальным вопросам развития этой индустрии в России, обсуждению ее места в мире, так и конкретным примерам применения и реализации суперкомпьютерных технологий. В рамках журнальной публикации невозможно осветить все темы форума. Мы отобрали несколько пленарных докладов, в которых, на наш взгляд, представлены основные тенденции развития и практического применения российских суперкомпьютеров.

**П**о сравнению с предыдущим годом НСКФ-2014 собрал меньше посетителей – 300 против 449 на НСКФ-2013. Однако на качество мероприятия это не повлияло. Докладов прозвучало больше, при этом в рамках форума впервые прошла конференция-сателлит "Посткремниевые вычисления", которой будет посвящена отдельная публикация.

Третий по счету НСКФ, окончательно закрепивший за собой статус регулярного мероприятия, претерпел важные организационные перемены. Для того чтобы координировать взаимодействие участников не только в дни работы форума, но и в остальное время, была создана автономная некоммерческая организация (АНО) "НСКФ". В ее обязанности, помимо организации мероприятия, входит управление интернет-порталом НСКФ, который играет роль связующего звена между участниками. Об этом в докладе, открывавшем форум, рассказал директор ИПС РАН, член-корреспондент РАН Сергей Михайлович Абрамов. Он особо подчеркнул нейтральность и общедоступность форума, участники которого – будь то организация или частное лицо – имеют равные права и возможности.

Одна из первых инициатив АНО "НСКФ" – учреждение ежегодной отраслевой премии НСКФ. Ее цель – усилить роль и причастность заинтересованных специалистов, готовых внести вклад в развитие суперкомпьютерной отрасли России. Премия будет символизировать заслуги, признанные не только индустрией, но и наукой, бизнес-сообществом. Выдвинуть кандидата на получение премии сможет любой участник форума, описав его заслуги в области суперкомпьютеров. Определен примерный механизм выбора лауреатов премии: окончательный список из 30 кандидатов формируется путем тайного голосования до начала НСКФ; три лауреата премии определяются во время проведения форума также тайным голосованием по методу Шульце, исключающим ошибки и неоднозначность результатов; лауреаты будут награждаться почетными знаками и денежными призами.

Традиционная часть выступления С.М.Абрамова – подведение итогов года. 2014-й, несмотря на печальные события и ухудшение политического и финансового климата в России, был весьма благоприятным для суперкомпьютерной отрасли. Чуда, конечно, не произошло, наша страна по-прежнему отстает от ведущих суперкомпьютерных держав.

Однако это отставание резко сократилось (рис.1). Доля России в "пироге" общей вычислительной мощности увеличилась с 0,7% в начале года до 1,6% в его конце (однако до рекордных 2,5% в 2010-м – еще далеко).

В список TOP500 2014 года вошло четыре (против двух в прошлом году) суперкомпьютера российской разработки: "Политехник РСК Торнадо" мощностью 658 Tflops, "Политехник РСК PetaStream" – 170 Tflops (обе машины находятся в СПбГТУ), "Ломоносов-2" – 320 Tflops ("Т-Платформы", установлен в МГУ), а также впервые попавшая в TOP500 суперЭВМ компании "Ниагара Компьютерс" GPU Blade cluster (289 Tflops) в ННГУ. Появление еще одного сильного игрока на российском рынке мощных вычислительных систем не может не радовать.

### 30 ЛЕТ ИПС РАН. Т-СИСТЕМА

Форум прошел под знаком 30-летия основателя и "дома" НСКФ – Института программных систем им. А.К.Айламазяна РАН. Истории института и его вкладу в развитие суперкомпьютерной отрасли была посвящена большая часть доклада С.М.Абрамова.

Институт программных систем РАН был создан в 1984 году. Последние два десятилетия существования СССР были весьма продуктивным периодом в истории отечественных суперкомпьютерных разработок: различные институты вели одновременно более десяти проектов в области многопроцессорных вычислений. Первой задачей ИПС РАН стала разработка ассемблера (как языка и системы программирования) и компилятора с языка С для мультипроцессора с динамической архитектурой ЕС-2704. В этой системе было реализовано несколько интересных решений, например, автоматическое динамическое распараллеливание пользовательских программ, устойчивость процесса выполнения программы к отказам части оборудования. ЕС-2704 содержал шесть интерфейсных, 12 коммутационных и 24 вычислительных процессорных модуля. Опыт работы с этой системой лег в основу последующих проектов, в том числе и создания Т-системы – среды программирования с автоматическим динамическим распараллеливанием задач.

Т-система стала первой важной вехой в истории отечественных суперкомпьютеров. Ее основные принципы были четко сформулированы в 1995 году. В качестве входного языка системы рассматривались диалекты известных языков, дополненных специальными конструкциями и функционально-ориентированными ограничениями.

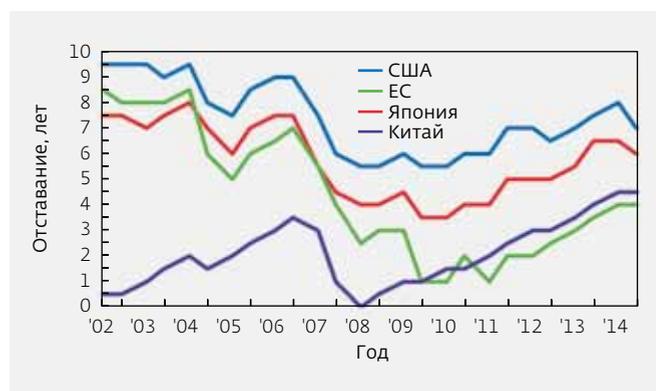


Рис.1. Отставание России по суммарной производительности суперкомпьютеров

Первая экспериментальная реализация Т-системы создавались в среде MS DOS, но параллельно велись работы по использованию ОС Linux и локальных сетей UNIX-станций. Была разработана сетевая компонента, обеспечивающая возможность распределенной загрузки задачи и внутризадачного обмена управляющими и информационными сообщениями.

В качестве аппаратной базы использовались различные сети из ПЭВМ – начиная с оригинальных сетей на базе ускоренных до 1 Мбит/с линий RS-232, собственных коммутирующих устройств для таких связей и заканчивая кластером на базе FastEthernet (100 Мбит/с).

Принципы распараллеливания задач исследовались на основных алгоритмах вычислительной математики. Было доказано, что специальные списковые структуры в Т-системе позволяют представлять характерные для таких задач данные и не приводят к существенной потере производительности системы.

Первая версия ядра Т-системы была опробована и отлажена на примере решения задачи построения реалистичных изображений виртуальных сцен методом трассировки лучей. Задание выполнялось в однопроцессорном режиме и локальной сети (Ethernet 10BASE-2) из четырех одинаковых однопроцессорных ПЭВМ с ОС Linux. Результаты счета на одном, двух, трех и четырех процессорах показали, что Т-система обеспечивает низкий (1,5–3%) уровень накладных расходов по сравнению с традиционной непараллельной реализацией, а рост производительности при увеличении количества вычислительных ядер близок к линейному. Первую рабочую реализацию Т-системы в кластере из четырех машин можно считать началом современного этапа суперкомпьютерных исследований в ИПС РАН.

1995 год был крайне продуктивным для института. Стоит упомянуть несколько важных разработок ИПС РАН. Одна из них – создание интерфейсной платы для ПЭВМ на основе транспьютера T 425. Применение разделяемой памяти обеспечило высокую скорость передачи данных по шине ISA – до 5 Мбайт/сек, что на порядок превосходило параметры всех существовавших в то время транспьютерных плат. Часть исследований в этой области была поддержана грантом INTAS, полученным вместе с итальянским филиалом компании Inmos (разработчика и производителя транспьютеров – больших интегральных схем, выступающих в роли "кирпичиков" для построения суперкомпьютерных систем) и с Университетом города Катанья.

В том же году был завершен перенос свободного компилятора GNU C Compiler на архитектуру транспьютеров семейств T 4, T 8 и T 9. Заключительная отладка и тестирование компилятора осуществлялись в удаленном режиме на установке GCell фирмы Parsytec в High Performance Computing Laboratory в Афинах (Греция). Использование Глобальной сети для проведения вычислений было большим шагом вперед, так как Интернет еще не был широко распространен в нашей стране. Тестирование показало, что по качеству генерируемого кода российский компилятор не уступает коммерческому компилятору ACE, входящему в состав ОС Parix. Этим результатам более чем десятилетней давности посвящен раздел в архиве Internet Parallel Computing Archive ([wotug.org/parallel](http://wotug.org/parallel)).

В 1998 году был введен в эксплуатацию кластер из 24 вычислительных узлов на базе процессоров Intel Pentium Pro 200 МГц и Pentium II 266 МГц производительностью 5,6 Gflops. В первые полгода эксплуатации кластера на примере десяти задач различной алгоритмической природы были подтверждены преимущества T-системы – почти линейная зависимость производительности от количества вычислительных ядер (до 24 возможных на данном кластере) и переносимость задач на T-языке на вычислительные узлы любой конфигурации без перекомпиляции и модификации. Было доказано, что T-система позволяет снизить затраты на разработку программ (автоматизация распараллеливания), увеличить глубину параллелизма и более полно использовать возможности аппаратной части мультипроцессора за счет распараллеливания в динамике.

### СУПЕРКОМПЬЮТЕРЫ СКИФ

T-система легла в основу проекта "СКИФ" – суперкомпьютерной программы, созданной совместно

с коллегами из Беларуси и ставшей еще одним важнейшим событием не только для института, но и всей суперкомпьютерной отрасли. Государственное финансирование программы началось в 2000 году, а к 2003-му было построено 16 опытных образцов и вычислительных установок рядов 1 и 2 семейства "СКИФ".

Параллельно совместно с МГУ разрабатывалось программное обеспечение – модифицированное ядро ОС LINUX под названием SKIF, пакеты параллельной файловой системы PVFS-SKIF и системы пакетной обработки задач OpenPBS-SKIF, мониторинговая система кластеров семейства "СКИФ" и т.д. Также были созданы две прикладные программы – система автоматизации проектирования химических реакторов (ИВВиИС, СПб) и система классификации потока текстов при помощи искусственного интеллекта (ИПС РАН).

В качестве базы метакластерной распределенной вычислительной сети, включающей системы "СКИФ" ИПС РАН, НИИ механики МГУ (Москва) и ОИПИ НАН Беларуси (Минск), была успешно опробована T-система.

Результаты программы "СКИФ" не заставили долго себя ждать. В ноябре 2003 года суперЭВМ "СКИФ К-500" заняла 405-е место в рейтинге TOP500, показав пиковую производительность 716,8 Gflops и реальную (на задачах Linpack) – 471,6 Gflops (65,79% пиковой). На момент создания "СКИФ К-500" был самой мощной вычислительной системой в России, СНГ и Восточной Европе. Ровно через год суперЭВМ "СКИФ К-1000" заняла уже 98-е место в рейтинге TOP500 с пиковой и реальной производительностью 2534 Gflops и 2032 Gflops (80,19% пиковой) соответственно. В рейтинге Top Cranch (поддержан агентством перспективных исследований МО США DARPA) суперкомпьютер "СКИФ К-1000" занял тогда первое место в мире при решении задачи расчета столкновения трех автомобилей.

### РАЗВИТИЕ T-СИСТЕМЫ И СОТРУДНИЧЕСТВО С MICROSOFT

Дальнейшим развитием T-системы стали ее версии GRACE (расширение языка C, 2002 год) и Open-TS с открытой архитектурой, реализованной в виде надстройки над языком C++. Open-TS эффективно поддерживает архитектуру SMP, асинхронный режим обменов внутри T-системы, механизм обеспечения отказоустойчивости, новую систему сбора статистики и механизм трассировки T-программ. В соответствии с новыми возможностями системы были модифицированы языковые средства T-системы – компилятор TG++, конвертор с языка T++ и транслятор с языка T-Fortran.

В 2004 году на Open TS обратила внимание корпорация Microsoft, предложив проверить эффективность использования OpenTS и библиотеки MPI для написания прикладных программ. В качестве "подопытных" были предложены два приложения на базе библиотеки MPI: Powray (пакет построения реалистичных трехмерных сцен методом трассировки лучей) и ALCMD (программа для расчетов в области молекулярной динамики). Эти приложения разрабатывались в течение многих лет усилиями сотен программистов – сторонников движения Open Source и оптимизированы настолько, насколько позволяет MPI-подход. В течение трех месяцев сотрудники ИПС РАН должны были переписать приложения с MPI на OpenTS. Результат считался бы положительным, если код T-системы получился более ясным и компактным. При этом допускалось снижение производительности до 30%. Тем самым корпорация Microsoft явно выражала свои предпочтения: повысить производительность труда программистов и создать более ясный, компактный и, как следствие, более надежный код (меньше строчек – меньше ошибок, больше ясности – меньше причин для ошибок).

Результаты работы оказались неожиданными даже для ее исполнителей. Команда из нескольких программистов в указанный срок переписала на OpenTS и отладила огромные программные комплексы. Полученный код был в 7-15 раз компактнее, чем его MPI-аналоги при сходной производительности. Корпорация Microsoft была удовлетворена столь выдающимися результатами и предложила институту контракт на перенос OpenTS из ОС Linux в Windows Compute Cluster Server. Проект был успешно реализован в 2006 году: ядро T-системы перенесено, транслятор с языка T++ интегрирован в Visual Studio, разработаны интерактивный инсталлятор, справочная подсистема и другие компоненты, традиционные для систем программирования под Windows. Результаты проекта были представлены на международной конференции и выставке Supercomputing'2006 (г. Тампа, Флорида, США, ноябрь 2006 года). В последующие годы (2007–2008) новые версии OpenTS создавались в кроссплатформенном варианте – один и тот же код OpenTS можно было собрать в дистрибутивы для Linux и Windows Compute Cluster Server.

Параллельно создавалось программное обеспечение для систем "СКИФ". Различными институтами были разработаны несколько прикладных программ, в том числе:

- MultiGen (ЧелГУ, создана в OpenTS) – система расчета биологической активности молекул

с учетом их конформационного многообразия, прогнозирования и проектирования в химии (лекарств и других соединений);

- пакет расчета аэромеханики подвижных плохообтекаемых тел (НИИ механики МГУ им. М.В.Ломоносова, создан в OpenTS);
- программная система моделирования широкополосных пространственно-временных радиолокационных сигналов;
- программная система поточечной обработки цветных и полутоновых видеоданных космических систем дистанционного зондирования (НИИ КС, созданы в OpenTS);
- модель профессора В.М.Лосева и другие метеорологические модели (ИПС РАН, Росгидромет);
- численные методы прогнозирования погоды, модели регионального прогноза погоды на 48 часов (ОИПИ НАН Беларуси, Республиканский гидрометеорологический центр);
- три прикладные системы искусственного интеллекта: АКТИС – классификация текстов по заданным в процессе обучения классам (глубокий анализ текста, высокая релевантность), INEX – извлечение знаний из неструктурированных текстов на естественном языке, заполнение заданной реляционной базы данных, MIRACLE (OpenTS) – инструмент для проектирования интеллектуальных систем (ИПС РАН);
- кардиологический комплекс на базе кластера "СКИФ" (ОИПИ НАН Беларуси, РНПЦ "Кардиология" и УП "НИИЭВМ");
- кардиологическая экспертная система реального времени "АДЕРТ-С" (ИВВиИС) и т.д.

### ПРОГРАММА "СКИФ-ГРИД". СУПЕРЭВМ "СКИФ" ТРЕТЬЕГО И ЧЕТВЕРТОГО ПОКОЛЕНИЙ

Летом 2007 года стартовала программа "Разработка и использование программно-аппаратных средств грид-технологий и перспективных высокопроизводительных (суперкомпьютерных) вычислительных систем семейства "СКИФ", головным исполнителем которой в России стал ИПС РАН. Помимо исследования, создания и внедрения грид-технологий, организации высокопроизводительных вычислительных сетей, программа предусматривала разработку суперкомпьютеров "СКИФ" третьего и четвертого поколений, а также реализацию комплексных средств защиты информации в проектируемых комплексах и подготовку кадров для суперкомпьютерной отрасли.

В рамках первого этапа исполнения программы была, в частности, создана распределенная вычислительная система "СКИФ-Полигон",

объединяющая суперЭВМ "СКИФ" в различных городах России, стран СНГ и научно-исследовательские центры, нуждающиеся в высокопроизводительных вычислениях. Суммарная пиковая производительность системы "СКИФ-Полигон" сегодня превышает 100 Tflops.

К марту 2008 года завершилось создание еще двух суперкомпьютеров "СКИФ" ряда 3: систем "СКИФ МГУ" (Linpack-производительность 47,17 Tflops, пиковая – 60 Tflops) и "СКИФ Урал" (Linpack-производительность 12,2 Tflops, пиковая – 16 Tflops). Обе суперЭВМ разработаны на базе одинаковых технических решений, предложенных в рамках первого этапа программы "СКИФ-грид". Впервые в России были применены 42-нм четырехъядерные процессоры Intel Xeon E5472 с частотой 3 ГГц. Серверы-лезвия (blade) отечественного производства отличались большей на 20% плотностью вычислительной мощности по сравнению с имевшимися аналогами. В июне 2008 года "СКИФ-Урал" вошла в Top500, заняв 282-е место, а "СКИФ-МГУ" с пиковой производительностью 60 Tflops заслужила 36-е место.

Современный флагман суперЭВМ "СКИФ" ряда 4 – система "СКИФ-Аврора", установленная в ЮУрГУ. Специально для реализации этого проекта сотрудники ИПС РАН учредили компанию "РСК СКИФ". Работы по созданию "СКИФ-Аврора" выполнялись совместно со специалистами итальянской компании Eurotech. Благодаря примененному в суперЭВМ водяному охлаждению удалось в полтора раза повысить эффективность использования электроэнергии и в два раза – плотность вычислительной мощности по сравнению с системами предыдущих поколений. Печатные платы и механические компоненты "СКИФ-Аврора" изготавливаются в России. Отсутствие в системе движущихся частей означает повышенную надежность узлов суперкомпьютера.

Вычислительные модули "СКИФ-Аврора" построены на базе процессоров Intel Xeon X5680 (шесть ядер с тактовой частотой 3,33 ГГц). Общий объем оперативной памяти составляет 3 ТБайт, системная сеть конфигурации "трехмерный тор" работает со скоростью 60 Гбит/с (максимальная задержка 1 мкс). Управляющая сеть – InfiniBand QDR, 40 Гбит/с с задержкой 2 мкс. В сентябре 2014 года "СКИФ-Аврора ЮУрГУ" с пиковой производительностью 117 Tflops занял 11-е место в рейтинге TOP500.

Результаты комплексной, хорошо продуманной программы "СКИФ", обеспечившей кумулятивный эффект для всей отрасли, говорят сами за себя. За несколько лет удалось создать суперкомпьютеры

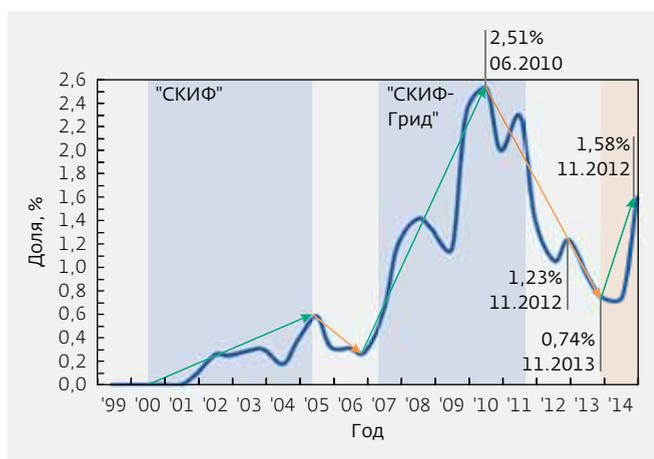


Рис.2. Доля российских суперкомпьютеров в мировой вычислительной мощности

мирового уровня, шесть машин "СКИФ" 22 раза попадали в TOP500, занимали места со второго по седьмое в списке самых мощных российских машин. Резкий рост доли России в вычислительной мощности суперкомпьютеров всего мира обусловлен именно развитием программы "СКИФ" (рис.2). Технические решения, разработанные в ее рамках, применяются в большинстве отечественных суперЭВМ, вошедших в TOP500 (см. таблицу).

Но самым важным результатом программы следует признать восстановление и создание кооперационных связей и организацию команды исполнителей, которой по плечу наиболее сложные задачи в области суперкомпьютерных технологий. Это хорошо иллюстрирует пример машины "К-1000", оказавшейся в 1994 году на 98-м месте в TOP500. В то время 100 первых мест занимали вычислительные системы из США, Японии и Китая, а страны Европы, имеющие собственные суперкомпьютеры, не могли похвастаться мощностями, достаточными для попадания в первую сотню рейтинга.

Нужно отметить, что на девять лет действия программы, в которой участвовало 27 организаций, было выделено всего 571 млн. рублей, а три года бюрократических процедур подписания документов буквально подкосили потенциал проекта "СКИФ-грид". Несмотря на это, "СКИФ" смело можно признать одной из самых успешных программ в России и Беларуси.

## СУПЕРКОМПЬЮТЕРНЫЙ ИНЖИНИРИНГ – НОВАЯ ПАРАДИГМА ПРОИЗВОДСТВА

Последняя четверть XX века прошла под знаком всеобщей компьютеризации промышленности. Разработки с помощью компьютера (CAD,



Computer-aided Design) стали одной из важнейших вех эволюции не только производства, но и цивилизации в целом. По мере развития концепция САД преобразовалась в инжиниринг с помощью компьютера (CAE) – понятие значительно более широкое, чем разработка, а появление доступных производству суперкомпьютеров привело к качественному изменению парадигмы CAE. О том, как суперкомпьютерный инжиниринг изменяет производство и насколько он важен для промышленности страны, рассказал директор отделения информационно-вычислительных ресурсов СПбГТУ, д.т.н., профессор Юрий Яковлевич Болдырев.

Под инжинирингом понимается интеллектуальная деятельность, в задачи которой входит проектирование, создание, сооружение, использование, переработка или утилизация в сфере конкретных технических задач. В такой широкой трактовке инжиниринг охватывает все отрасли человеческой деятельности – от аэрокосмической промышленности до производства продуктов питания.

В последнее десятилетие в развитии компьютерного инжиниринга появились новые тенденции, которые позволили существенно расширить и углубить области его применения. Вкупе с одно-временным внедрением передовых производственных технологий (AMT – Advanced Manufacturing Technologies) и достижений науки о материалах, направленной на исследования новых керамик, композитов и т.д., можно говорить о становлении принципиально нового направления в фундаментальных основах инженерного знания – суперкомпьютерного инжиниринга.

Каковы основные отличительные черты суперкомпьютерного инжиниринга? В первую очередь, это, конечно, стремительно растущие и практически неисчерпаемые вычислительные мощности. Если в начале нынешнего века вычислительная производительность в несколько терафлопс была рекордной, то в наши дни нормой для развитых стран становятся петафлопсные системы. Эта тенденция открывает вторую, гораздо более важную особенность суперкомпьютерного

Суперкомпьютеры российского производства в TOP500

Дата вхождения в TOP500	Место	Производительность*, Tflops	Название	Создатель
6/2002	64	0,734	MBC 1000M	ФГУП "НИИ Квант", МСЦ РАН, ИПМ им. М.В.Келдыша РАН
11/2003	405	0,423	"СКИФ" К-500	СКИФ-кооперация
11/2004	98	2,03	"СКИФ" К-1000	
6/2007	105	9,01	"СКИФ Cyberia"	
6/2008	282	12,2	"СКИФ Урал"	
6/2008	36	47,1	"СКИФ МГУ Чебышев"	
11/2009	452	21,8	"СКИФ-Аврора", ЮУрГУ	
11/2009	12	350	"Ломоносов", МГУ	ГК "Т-Платформы"
11/2012	59	376	MBC-10 "Tornado" МСЦ РАН	ГК "РСК"
11/2012	170	147	Polytechnic RSC Tornado, ЮУрГУ	
6/2014	129	320	"Ломоносов-2" (A class, МГУ)	ГК "Т-Платформы"
11/2014	81	658	Polytechnic RSC Tornado, СПбГПУ	ГК "РСК"
11/2014	189	289	"Лобачевский" GPU Blade Cluster, ННГУ	Niagara Computers, Supermicro
11/2014	390	170	RSC PetaStream СПбГПУ	ГК "РСК"

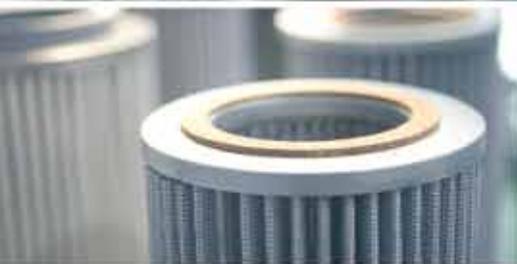
\* на момент вхождения в TOP500

10-я Международная выставка  
вакуумной техники, материалов и технологий

14–16 апреля 2015 года  
Москва, КВЦ «Сокольники»

единственная в России и СНГ

выставка вакуумной отрасли



забронируйте стенд на сайте

[www.vacuumtechexpo.com](http://www.vacuumtechexpo.com)

- Вакуумная техника и компоненты
- Средства получения и измерения вакуума
- Криогенное оборудование и комплектующие

- Течеискатели и аналитическое оборудование, работающее в вакуумной среде
- Оборудование для нанесения функциональных покрытий

Организатор:



Тел.: +7 (495) 935 81 00

E-mail: [vacuumtechexpo@ite-expo.ru](mailto:vacuumtechexpo@ite-expo.ru)

При поддержке:



НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ  
ИНСТИТУТ ВАКУУМНОЙ ТЕХНИКИ  
ИМ. С.А. ВЕКШИНСКОГО

Генеральный спонсор:



инжиниринга – возможность решения междисциплинарных задач, которые максимально точно моделируют реальный физический мир. Процессы, происходящие в реальном мире, зависят от множества факторов, принадлежащих различным областям науки. Например, для описания важного для моделирования тепловых двигателей процесса горения необходимо аккумулировать знания из разных разделов физики, химии, газодинамики, учитывать уровень излучения и т.д. Это и есть междисциплинарность – крайне наукоемкая парадигма, требующая колоссальных вычислительных мощностей. И чем "тоньше" модель процесса, то есть чем больше факторов учитывает, тем ближе она к реальности.

По мере усложнения промышленных разработок востребованы принципиально новые, интеллектуальные методы и технологии. При этом требуется создавать конкурентоспособную продукцию в крайне сжатые сроки – а это значит, что производство должно быть оптимизировано. Все это способствует более широкому внедрению компьютерного инжиниринга, поскольку именно такие задачи успешно решаются на основе методов математического моделирования.

На базе суперкомпьютерного инжиниринга, новых технологий и наук сформировалось несколько направлений математического моделирования, которые различаются целями выполняемых задач – моделирование изделий и систем, поведения материалов, оптимизация изделий и производственных процессов. Комплекс таких технологий формирует концепцию цифрового производства, на котором вместо реальных моделей создается виртуальный "цифровой прототип" изделия. Появляется возможность исключить из процесса разработки изделия процедуру создания дорогостоящих натуральных прототипов и моделировать любые характеристики изделия путем расчетов на суперкомпьютерах.

Наиболее ярким примером реализации подобного подхода служат автомобильные краш-тесты. Не связанные напрямую с проектированием автомобилей, они служат важнейшим инструментом для определения слабых с точки зрения безопасности мест в конструкции автомобиля. Конечно, адекватность компьютерных краш-тестов зависит непосредственно от точности нелинейных математических моделей, быстро протекающих процессов в средах с крайне сложной структурой, в том числе имеющих полости, содержащие жидкости.

В докладе Ю.Я.Болдырев представил два главных тезиса: универсальность суперкомпьютерного

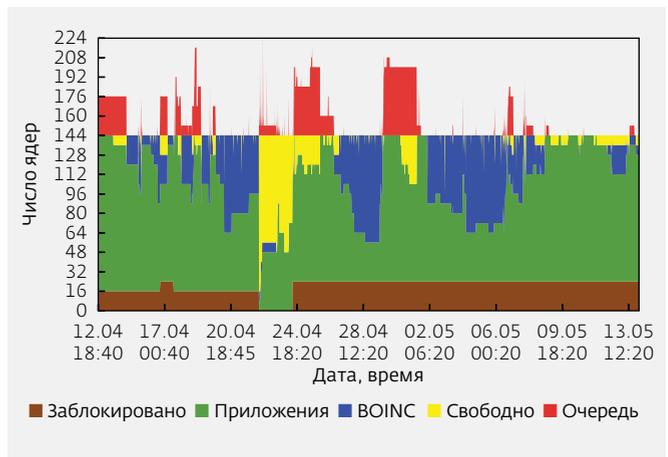


Рис.3. Уровень загрузки кластера Blackford IDCiTY СО РАН с системой CluBORun в период с 12 апреля по 13 мая 2013 года

инжиниринга и возможность решения любой корректно сформулированной задачи. Второй тезис может показаться излишне смелым заявлением, но это не так. Сегодня успешно решаются задачи, постановка которых 15–20 лет казалась невозможной. И технологии суперкомпьютерного инжиниринга продолжают совершенствоваться.

Тем не менее нельзя утверждать, что решение современных инженерных задач не вызывает трудностей. Некоторые алгоритмы расчета физических процессов все еще слишком сложны для того, чтобы обеспечить получение результата в приемлемое время на современной аппаратной базе. При разработке математических моделей и переходе от непрерывных моделей к дискретным крайне важно проверять их корректность, чему не всегда

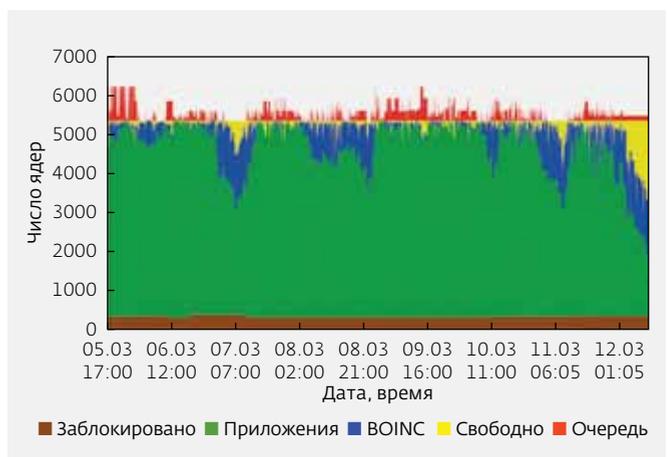


Рис.4. Загрузка кластера MBC-100K в период с 5 по 12 марта 2014 года

12+

реклама



С В Я З Ъ  
Э К С П О  
К О М М  
2 0 1 5

12–15 мая

 ЭКСПОЦЕНТР

Организаторы:

- ЗАО «Экспоцентр»
- Фирма «И.Джей.Краузе энд Ассосиэйтс, Инк.» (США)

При поддержке:

- Министерства связи и массовых коммуникаций РФ
- Министерства промышленности и торговли РФ
- Федерального агентства связи (Россвязь)
- Правительства Москвы

Под патронатом

Торгово-промышленной палаты РФ

27-я международная  
выставка

телекоммуникационного  
оборудования,  
систем управления,  
информационных  
технологий и услуг связи

Россия, Москва, ЦВК «Экспоцентр»

[www.sviaz-expocomm.ru](http://www.sviaz-expocomm.ru)



придается должное значение. Между тем некорректные расчеты могут привести к убыткам и человеческим жертвам.

Нужно понимать также, что успешный суперкомпьютерный инжиниринг – это не только математические модели и вычислительные системы. Известный лозунг гласит: кадры решают все. Именно грамотные инженеры и руководители определяют успех суперкомпьютерных технологий. В английском языке прижился термин *brainware* – "мозговое обеспечение", который подчеркивает важность интеллектуальных ресурсов, ставит их в один ряд с компонентами компьютеров – *hardware* (аппаратное обеспечение) и *software* (программы). А грамотное планирование и руководство (как на уровне производства, так и вышестоящее) культивируют на предприятиях атмосферу востребованности современных технологий для победы в конкурентной борьбе.

Не менее важный фактор успешного внедрения суперкомпьютерного инжиниринга – образование. Источники *brainware* – вузы и научно-технические центры. Для того чтобы подогревать интерес учащихся к суперкомпьютерному инжинирингу, помогать им получать нужные для этого знания, необходимо обеспечить тесное взаимодействие с учебными и научными учреждениями. Это крайне важно и в глобальном плане: падение уровня физико-математического образования в школах, преподавание фундаментальных наук в вузах в усеченном варианте неизбежно приводит к ослаблению суперкомпьютерной отрасли и инжиниринга в частности. Уже сегодня многие ведущие российские предприятия испытывают нехватку кадров в области математического моделирования. Один из способов решения проблемы – создание в ведущих региональных вузах центров суперкомпьютерного инжиниринга, которые координировали бы взаимодействие с промышленностью. Сегодня в России существует всего один такой центр на базе СПбГТУ.

Мировой опыт показывает, что инициатором поддержки суперкомпьютерного инжиниринга должно быть государство. Например, в США содействие развитию этого направления оказывают министерства обороны и энергетики. В России государство только начинает принимать участие в организации поддержки суперкомпьютерным инжиниринговым технологиям: осенью 2014 года при Минпромторге был сформирован технический комитет "Математическое моделирование и высокопроизводительные вычисления", в состав которого вошли опытные специалисты в этой области.

## CluBORun – РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛЕНИЯ НА РЕСУРСАХ СУПЕРКОМПЬЮТЕРНЫХ КЛАСТЕРОВ

Для современных технологий высокопроизводительных вычислений актуальны две проблемы. Первая – задачи, решение которых и на очень мощном суперкомпьютере может занять продолжительное время (месяцы и даже годы). Получить на такой срок суперкомпьютер для выполнения одной задачи, как правило, нереально – за исключением ситуаций, когда суперкомпьютер создан специально "под конкретную проблему" (например, в ядерных центрах).

Вторая проблема, возникающая при работе с современными суперкомпьютерами, состоит в их неравномерной загрузке. В один промежуток времени суперкомпьютер может быть загружен более чем на 90%, а в другие – менее чем наполовину. Беда в том, что даже при значительной недогрузке суперкомпьютер продолжает потреблять большой объем электроэнергии. Мониторинг в течение года показал, что загрузка суперкомпьютеров составляет в среднем всего около 50% и даже меньше. Лишь некоторые кластеры загружены на 70%.

Ситуация, на первый взгляд, парадоксальная: с одной стороны, существуют востребованные практикой задачи, для решения которых не подходят современные политики использования суперкомпьютеров, а с другой – значительную часть времени суперкомпьютеры не загружены на полную мощность, что приводит к удорожанию вычислений.

В качестве одного из решений этих проблем была предложена концепция многозадачной грид-системы с гибким распределением свободных ресурсов на примере проекта CluBORun (Cluster for BOINC Run), разработанного в Институте динамики систем и теории управления (ИДСиТУ) СО РАН (Иркутск). О нем рассказал директор ИДСиТУ СО РАН, академик Игорь Вячеславович Бычков.

В основе CluBORun лежит хорошо известная концепция добровольных распределенных вычислений (ДРВ). В рамках ДРВ вычислительные ресурсы, как правило, собственные персональные компьютеры, предоставляются частными лицами (добровольцами-энтузиастами). Для проектов ДРВ подходят задачи, которые могут быть разбиты на подзадачи меньшей размерности, решаемые независимо. Сложность элементарных задач должна быть такой, чтобы их решение на маломощном компьютере занимало не более нескольких часов.

Наиболее известная платформа для ДРВ – BOINC, функционирующая на основе следующих принципов:

- управление проектом осуществляет центр администрирования на базе выделенного сервера. В его обязанности входит создание клиентских приложений, генерация пакетов элементарных заданий, их рассылка участникам, сбор и обработка присланных ими результатов. Один центр может управлять несколькими VOINC-проектами;
- на компьютерах – узлах сети (хостах) устанавливается программа VOINC-менеджера, которая принимает пакеты задач, следит за их решением и распределяет ресурсы узла так, чтобы решение задач не мешало работе пользователя, например, система загружается только в моменты простоя;
- на одном хосте возможно решение нескольких задач, их приоритет задается программой-менеджером.

Указанные принципы и легли в основу концепции грид-системы с гибким распределением ресурсов, только вместо персональных компьютеров роль хостов играют вычислительные узлы кластеров. Точно так же, как и в VOINC-сети, к вычислениям элементарных задач привлекаются не занятые основной работой кластеры. Важно, что грид-система обращается к кластеру только с правами стандартного пользователя, не конфликтуя с политиками администрирования кластера.

Механизм работы с кластером отличается от такового для персональных компьютеров. В каждом кластере работает система управления очередью заданий, которая распределяет свободные ресурсы между задачами пользователей. Для использования свободных ресурсов кластера необходимо отслеживать очередь и занимать доступные ресурсы под видом обычного пользователя. Эти обязанности выполняет автоматическая система, названная СИСР (система использования свободных ресурсов).

От СИСР требуется определенная деликатность. Даже при наличии значительного количества свободных ресурсов она не должна занимать их надолго, иначе у зашедшего в кластер пользователя возникнет "эффект испуга". Пользователь должен быть уверен в том, что необходимые ресурсы высвободятся в ближайшее время. Если СИСР запустила задания, например, на несколько часов, а в очереди появляются другие пользователи, система должна приостанавливать свои задания, запоминая промежуточные результаты вычислений.

Цель проекта CluBORun – запуск клиентских приложений VOINC на вычислительных кластерах посредством VOINC-менеджера. Но поскольку

ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО

# ЭЛЕКОНД





## КОНДЕНСАТОРЫ

разработка и производство

**оксидно-электролитические  
алюминиевые**  
K50-15, K50-17, K50-27, K50-37, K50-68,  
K50-74, K50-76, K50-77, K50-80, K50-81,  
K50-83, K50-84, K50-85, K50-86, K50-87,  
K50-88, K50-89, K50-90, K50-91, K50-92

**объемно-пористые танталовые**  
K52-1, K52-1M, K52-1BM, K52-1B, K52-9,  
K52-11, K52-17, K52-18, K52-19, K52-20, K52-21

**оксидно-полупроводниковые  
танталовые**  
K53-1A, K53-7, K53-65(чип), K53-66, K53-68(чип),  
K53-71(чип), K53-72(чип), K53-74(чип)

**оксидно-полупроводниковые  
ниобиевые**  
K53-4, K53-52, K53-60

Россия, 427968, Удмуртская республика,  
г. Сарапул, ул. Калинина, 3  
Тел.: (34147) 2-99-53, 2-99-89, 2-99-77  
Факс: (34147) 4-32-48, 4-27-53  
e-mail: [elecond-market@mail.ru](mailto:elecond-market@mail.ru)  
<http://www.elecond.ru>

Система менеджмента качества сертифицирована на соответствие требованиям ГОСТ Р ИСО 9001-2008

стандартные задачи выполняются на кластере в виде MPI-приложений, а BOINC-менеджер таковым не является, была создана программа-посредник между интерфейсом MPI и BOINC-менеджером, запускаемая стандартным образом. СИСП CluBORun организована в виде двух скриптов, один из которых отвечает за мониторинг ресурсов кластера, а другой – за запуск на свободных узлах программы-посредника.

На текущем этапе проекта CluBORun используется для усиления вычислительной мощности проектов ДРВ SAT@home (решение дискретных задач) и ОПТИМА@home (реализация численных алгоритмов оптимизации). Следует отметить, что идея привлечения кластеров в проекты добровольных вычислений весьма плодотворна, поскольку ресурсы кластера надежнее ресурсов ПК как в плане предсказуемости времени решения соответствующих задач, так и достоверности результатов вычислений.

Основная техническая трудность при подключении к CluBORun-грид новых кластеров заключается в том, что на разных кластерах используются различные системы управления очередью заданий. Соответственно, перечисленные выше алгоритмы взаимодействия с очередью приходится каждый раз переписывать с учетом возникающих нюансов. Первая версия CluBORun-грид была создана для работы с системами управления очередями заданий кластеров Cleo (разработка НИВЦ МГУ) и СУПЗ (разработка МСЦ РАН). Эта же версия CluBORun-грид была запущена в кластере Blackford ИДСиТУ СО РАН со 144 ядрами. На рис.3 синим цветом выделены BOINC-расчеты, зеленым – расчеты, выполняемые другими пользователями кластера, желтым – свободные ресурсы, красным – затребованные другими пользователями в очереди заданий. Оранжевым цветом отмечены заблокированные ресурсы, то есть ядра, доступ к которым был закрыт политикой администрирования. В этом эксперименте кластер использовался в связке с проектом добровольных вычислений SAT@home.

С 22 по 27 апреля 2013 года на работу CluBORun накладывались жесткие ограничения по числу свободных узлов, из-за чего более половины ресурсов кластера простаивало. В следующий промежуток времени ограничения смягчались либо снимались, и все незаблокированные узлы кластера задействовались в обычных либо в BOINC-расчетах. Также на рис.3 хорошо видно, как CluBORun освобождает от BOINC-расчетов узлы при образовании очереди заданий других

пользователей: при появлении новых задач останавливалось ровно столько BOINC-расчетов, сколько требуется для запуска задач сторонних пользователей.

В декабре 2013 года CluBORun-грид была запущена в кластере МВС-100К с системой СУПЗ. Особенности кластера привели к появлению второй, а затем третьей версии CluBORun-грида с новыми алгоритмами обработки информации об очереди заданий.

Версия CluBORun-грид, развернутая в МВС-100к, позволяет использовать значительную часть ядер, не занятых обычными расчетными задачами (рис.4).

С помощью CluBORun в проекте SAT@Home были найдены 17 новых пар ортогональных диагональных латинских квадратов порядка 10 (до этого было известно три такие пары). Также были решены три ослабленные задачи криптоанализа генератора ключевого потока Vivium. Порой вклад кластера МВС-100К в производительность SAT@home достигал 40%! Помимо SAT@home, CluBORun-грид используется в связке с проектом ДРВ ОПТИМА@home.

В заключение рассказа о проекте CluBORun И.В.Бычков указал на его отличия от похожих грид-решений. Главное преимущество CluBORun – инструменты для отслеживания очереди заданий кластера и гибкого выделения его ресурсов. Существующие системы для объединения BOINC-систем и вычислительных кластеров (BNB-Grid, идеи которой легли в основу CluBORun, 3G-Bridge) таких возможностей не имеют. Немаловажно также, что исходный код CluBORun открыт и выложен в свободный доступ.

Предложенный подход может быть применен не только для добавления вычислительных ресурсов кластеров в проекты ДРВ. Перспективным представляется создание сервисных грид-систем, ориентированных на использование свободных ресурсов вычислительных кластеров, входящих в эти системы. Для управления такой грид-системой может быть задействован стандартный BOINC-сервер без функции подключения к нему частных лиц.

Возможно, в более далекой перспективе описанные в статье технологии могут быть использованы для создания многозадачной грид-системы с широким спектром решаемых задач и с привлечением свободных ресурсов большого числа географически удаленных друг от друга суперкомпьютеров.

Продолжение следует ●