

Высокопроизводительные вычислительные ресурсы Института динамики систем и теории управления СО РАН: Текущее состояние, возможности и перспективы развития*

И. В. Бычков, Г. А. Опарин, А. П. Новопашин, А. Г. Феокистов,
А. С. Корсуков, И. А. Сидоров

Институт динамики систем и теории управления СО РАН, Иркутск, Россия
e-mail: bychkov@icc.ru, oparin@icc.ru, apn@icc.ru, agf@icc.ru,
alexask@mail.ru, ivan.sidorov@icc.ru

Представлен опыт, накопленный в Институте динамики систем и теории управления СО РАН в области организации высокопроизводительных параллельных и распределенных вычислений.

Ключевые слова: высокопроизводительные вычисления, параллельные и распределенные системы, инструментальные средства, GRID.

Введение

В последнее десятилетие мировое IT-сообщество претерпевает значительные перемены, обусловленные стремительным ростом суперкомпьютерных мощностей, тотальной интеграцией вычислительных и коммуникационных систем, развитием GRID-систем различного назначения. Тем не менее в условиях быстро развивающейся компьютерной базы параллелизм, реализованный практически на всех уровнях вычислительной аппаратуры, используется далеко не полностью. Отсутствуют развитые языки, технологии и инструментальные средства, ориентированные на прикладных специалистов и обеспечивающие адекватное отображение задач на конкретную программно-аппаратную платформу. Крайне мало внимания уделяется построению прикладных программных систем на основе повторного использования программных блоков, особенно большого количества таких блоков. Эффективность создаваемых приложений до сих пор во многом зависит от степени владения низкоуровневым системным программным обеспечением (MPI, OpenMP, Globus Toolkit и т. д.), что доступно узкому кругу программистов высокой квалификации. Многие распределенные прикладные программные комплексы разрабатываются без соблюдения требований к переносимости, совместимости и адаптируемости к динамически изменяющейся аппаратной среде. Часто отсутствуют характеристики их функционирования в соответствии с требуемыми показателями качества.

Анализ мировых тенденций в области автоматизации решения прикладных задач в параллельных и распределенных вычислительных средах позволяет утверждать, что

*Работа выполнена при финансовой поддержке программы фундаментальных исследований Президиума РАН № 13 (проект СО РАН № 3).

© ИВТ СО РАН, 2010.

решение многих проблем непосредственно связано с интеллектуализацией так называемого промежуточного программного обеспечения (ПО), позволяющего динамически интегрировать распределенные гетерогенные ресурсы в виртуальную исполнительную среду и предоставляющего возможности для прозрачного применения этой среды.

Наши исследования [1–15] показывают, что комплексное использование методов концептуального крупноблочного программирования, мультиагентных технологий, логических методов синтеза и конструирования программ, технологий баз знаний и других средств искусственного интеллекта является ключевым фактором в создании удобных информационно-вычислительных сред для прикладных специалистов, нацеленных на эффективное применение высокопроизводительных ресурсов без погружения в особенности низкоуровневого параллельного и/или распределенного программирования решаемой задачи. В таких средах построение параллельной (распределенной) крупноблочной программы на основе библиотеки специфицированных прикладных модулей выполняется автоматически по целевому содержательному запросу.

В данной статье представлен опыт, накопленный в Институте динамики систем и теории управления (ИДСТУ) СО РАН в области организации высокопроизводительных параллельных и распределенных вычислений. В частности, описана материально-техническая база для супервычислений, приведены классы и примеры решаемых ресурсоемких научно-технических задач. Рассмотрены возможности разработанных в институте высокоуровневых средств автоматизации параллельного и распределенного программирования прикладных задач, а также инструментальных средств и технологий организации GRID-систем различного уровня. Обозначены направления работы ИДСТУ по подготовке специалистов (в том числе высшей квалификации) по параллельным и распределенным вычислениям и GRID-технологиям.

1. Инфраструктура для высокопроизводительных вычислений

С 2005 г. при ИДСТУ СО РАН функционирует Суперкомпьютерный центр (СКЦ) [16, 17], предоставляющий высокопроизводительные вычислительные ресурсы и системы хранения данных в коллективное пользование учреждениям Иркутского научно-образовательного комплекса для проведения наукоемких вычислений. Основу такой инфраструктуры составляют вычислительные установки, в которых реализованы наиболее востребованные на сегодня архитектурные решения — кластеры выделенных серверных узлов и серверы на базе графических ускорителей.

Вычислительный кластер МВС-1000/16 ИДСТУ СО РАН — разработан Институтом прикладной математики им. М.В. Келдыша РАН и ФГУП НИИ “Квант” (Москва), введен в эксплуатацию в феврале 2005 г. (первая очередь), спустя год модернизирован силами ИДСТУ до представленной в настоящее время конфигурации. Кластер включает 16 вычислительных узлов (32 процессора) суммарной пиковой производительностью 170 GFlops. В качестве системной сети используется Myrinet 2000, управляющей — Gigabit Ethernet. Машина поддерживается в состоянии максимальной работоспособности, однако по мере ее морального устаревания и с введением в строй более мощного вычислителя используется все менее интенсивно — главным образом для запуска адаптированных ранее приложений, а также в учебных целях.

Вычислительный кластер “Blackford” — основной вычислительный ресурс СКЦ, собственная разработка ИДСТУ СО РАН, введен в эксплуатацию в феврале 2007 г., в январе следующего года модернизирован, что позволило увеличить пиковую произво-

дительность до 1.5 TFlops и добиться реальной производительности на тестах LINPACK порядка триллиона операций с плавающей точкой в секунду. Кластер построен на базе серверных платформ Intel S5000PAL (Bensley), четырехъядерных процессоров Intel Quad-Core Xeon E5345 2.33 GHz (Clovertown), интерконнекта Gigabit Ethernet. Вычислитель содержит 20 вычислительных узлов, 40 процессоров (160 процессорных ядер). В сентябре 2008 г. “Blackford” вошел в девятую редакцию рейтинга наиболее мощных суперкомпьютерных систем СНГ и занял в нем 41-е место.

Оба кластера работают под управлением операционной системы Gentoo Linux и снабжены средствами создания, отладки и запуска параллельных программ, включающими компиляторы C/C++ и Fortran, коммуникационные интерфейсы (MPI и PVM), параллельные предметные библиотеки (ScaLAPACK, MKL, ATLAS, FFT, PETSc) и системы параллельного программирования (DVM и HPF-adaptor); средствами мониторинга и управления процессом прохождения параллельных задач; прикладным ПО (пакет GAMESS для квантово-химических расчетов, параллельный SAT-солвер PMSat, пакеты Lamarck/Migrate, Parallel MrBayes для популяционно-генетических и филогенетических расчетов, пакет ILOG CPLEX для решения задач линейного и целочисленного программирования и др.).

С целью освоения перспективных параллельных архитектур, к числу которых относится программно-аппаратная платформа CUDA (Compute Unified Device Architecture), реализованная в некоторых сериях графических ускорителей (Graphics Processing Unit, GPU) компании NVIDIA и открывающая новые возможности для решения ряда ресурсоемких научно-исследовательских задач дискретной математики, молекулярной динамики, вычислительной биологии, криптографии и других областей, в конце 2009 г. в СКЦ введен в эксплуатацию *вычислительный сервер Supermicro SuperServer 7046GT-TRF с четырьмя GPU NVIDIA Tesla C1060*. Общее количество ядер в графических процессорах сервера — 960, их суммарная пиковая производительность — 3.73 TFlops для операций с одинарной точностью и 312 GFlops для операций с двойной точностью. Операционная система — Windows Server 2008. Программное обеспечение GPU-сервера включает также средства разработки (CUDA Toolkit, CUDA Developer SDK, MS Visual Studio Pro 2008), средства для выполнения математических расчетов (пакет Mathworks Matlab с надстройкой Jacket, позволяющей запускать стандартный код Matlab на графическом процессоре), пакеты для решения задач молекулярной динамики NAMD/CUDA, HOOMD/CUDA и др.

Для хранения больших массивов предметных данных в распоряжении пользователей СКЦ находится современная высокопроизводительная масштабируемая система хранения *T-plaforms ReadyStorage 3994* объемом 62 ТБ, реализованная в соответствии с архитектурой SAN (Storage Area Network). Система хранения данных связана с вычислителями через оптические каналы связи с пропускной способностью 4 Gbps.

Учитывая постоянный рост требований к производительности вычислительных установок коллективного пользования со стороны прикладных задач, а также увеличение количества и размерности таких задач (см. раздел 2), в ИДСТУ СО РАН ведутся проектные работы по созданию вычислительной установки нового поколения, на порядок превосходящей по производительности ныне действующие. Проект предполагает создание кластерного вычислительного комплекса (КВК) пиковой мощностью не менее 20 TFlops на базе многоядерных процессоров, blade-серверов, высокоскоростного интерконнекта и системы хранения данных с параллельной файловой системой. Для обеспечения необходимых условий для функционирования КВК предусмотрено развер-

тывание систем бесперебойного электропитания, охлаждения, автоматического пожаротушения, мониторинга и безопасности. К настоящему времени разработано техническое задание на выполнение опытно-конструкторских работ, проведена реконструкция машинного зала ИДСТУ для размещения КВК.

2. Классы и примеры решаемых задач

С использованием указанного выше оборудования в СКЦ решается широкий спектр задач, относящихся к следующим научно-исследовательским областям:

- геномный анализ,
- филогенетика,
- физика высоких энергий,
- физика твердого тела,
- квантовая химия,
- криптоанализ,
- дискретная математика,
- оптимальное управление динамическими системами.

Приведем примеры решаемых задач.

1. Анализ метагеномных данных методом таксономической классификации нуклеотидных последовательностей.

2. Исследование популяционной структуры байкальских сигаев путем анализа полиморфизма микросателлитных локусов [18].

3. Анализ ДНК последовательностей динофлагеллят (групп микроорганизмов, составляющих существенную часть сообществ морских и пресноводных одноклеточных) [19].

4. Исследование фазовой диаграммы квантовой хромодинамики и свойств адронов в горячей и плотной среде [20].

5. Серия расчетов, связанных с исследованием точечных дефектов в кристаллах щелочных и щелочно-земельных фторидов методами квантовой химии [21, 22].

6. Расчеты, относящиеся к спектроскопии колебательно-вращательных состояний молекул [23].

7. Разработка новых скоростных методов и параллельных алгоритмов логического поиска в задачах обращения дискретных функций. Решение задачи криптоанализа систем поточного шифрования [24–26].

8. Разработка методов и средств параллельного решения дискретных задач в булевых ограничениях [27, 28].

На решение перечисленных задач затрачивается более 90 % заказанного машинного времени кластеров СКЦ.

Остановимся более детально на тех задачах, которые требуют для своего решения вычислительных кластеров и/или GRID-систем повышенной (десятки TFlops) и сверхвысокой (сотни TFlops) производительности.

Построение фазовой диаграммы квантовой хромодинамики в рамках нелокальной кварковой модели (Раджабов А.Е., ИДСТУ СО РАН).

В связи с подготовкой и проведением экспериментов на ускорителях частиц ЛНС (ЦЕРН), FAIR (Дармштадт), NICA (Дубна) отмечается повышенный интерес к построению теоретических моделей, описывающих состояния вещества при экстремально высоких температурах и плотностях. Ожидается, что этот новый тип материи может быть

обнаружен в данных экспериментах. Теоретическое описание новых состояний вещества представляет собой задачу огромной сложности и требует привлечения не только новых теоретических методов и подходов, но и самой современной вычислительной техники. Последнее связано с тем, что предложенные модели включают вычисление сложных многопетлевых интегралов и бесконечных сумм по дискретным частотам энергии, а это требует годы расчетов на обычных персональных компьютерах.

Использование вычислительного кластера “Blackford” для расчетов свойств сильно-взаимодействующей материи при конечной температуре уже позволило получить ряд новых интересных результатов в описании адронной фазы. Однако дальнейшее исследование моделей, позволяющих учитывать не только конечнотемпературные эффекты, но и эффекты конечной барионной плотности, требует использования еще более мощной вычислительной техники. Это позволит быть на переднем крае мировых научных исследований в данной бурно развивающейся области физики высоких энергий.

Расчет свойств твердотельных материалов (Раджабов Е.А., Мысовский А.С., Мясникова А.С. и др., Институт геохимии им. А.П. Виноградова СО РАН).

С использованием кластерных вычислительных установок ИДСТУ СО РАН проводятся исследования точечных дефектов в кристаллах щелочных и щелочно-земельных фторидов с применением методов квантовой химии. К таким задачам, в частности, относятся:

— исследование сцинтилляционных свойств кристаллов щелочно-земельных фторидов с примесями Се и Cd, изучение оптических и электронных характеристик примесных центров;

— моделирование переходов с переносом заряда в тех же кристаллах с примесями Eu, Yb и других редких земель;

— установление пространственных конфигураций, оптических и магниторезонансных свойств фотохромных центров в тех же кристаллах с примесями Y, La и Lu;

— проверка существования слабонепосредственных конфигураций автолокализованных экситонов в кристаллах CaF_2 , SrF_2 , BaF_2 , уточнение конфигураций околопримесных экситонов и механизмов их образования и распада;

— исследование роли примесного иона Cu^+ в процессах запасаения энергии ионизирующих излучений и термостимулированной люминесценции в кристаллах фторида лития.

Все расчеты выполняются достаточно точными и при этом хорошо масштабируемыми неэмпирическими методами, такими как функционал плотности ВЗЛур, методы связанных кластеров, реже — традиционный метод Хартри—Фока и конфигурационное взаимодействие. Эксперименты требуют весьма значительных вычислительных ресурсов и будут тем успешнее, чем большие ресурсы окажутся в распоряжении исследователей.

Параллельные технологии в задачах обращения дискретных функций (Семенов А.А., Заикин О.С., Беспалов Д.В., Игнатьев А.С., ИДСТУ СО РАН).

Разрабатываются алгоритмы решения задач о пропозициональной выполнимости (так называемых SAT-задач) с последующей апробацией на задачах криптоанализа различных систем поточного шифрования. За последние два года создана параллельная технология решения SAT-задач, которая реализована в программном комплексе D-SAT для кластерных вычислительных систем. Данная технология показала свою эффективность при решении целого ряда задач криптоанализа генераторов ключевого потока (Геффе, Вольфрама, суммирующий, пороговый, Гиффорда, A5/1). Особый интерес

представляют результаты по параллельному криптоанализу генератора A5/1. Прогноз ресурсов, необходимых для полного решения задачи криптоанализа данного генератора, требует менее суток работы кластера, имеющего производительность 60 TFlops. Однако даже располагая менее мощными ресурсами, удается получать интересные результаты. Так, менее чем за сутки работы кластера “Blackford” для генератора A5/1 были найдены различные инициализирующие последовательности, порождающие один и тот же ключевой поток (коллизии). Созданная в ИДСТУ параллельная технология решения SAT-задач базируется на концепции крупноблочного параллелизма, и ее эффективность напрямую зависит от числа вычислительных единиц в используемой системе. Доступность “больших” кластеров и вычислительных GRID должна способствовать развитию этой и других (находящихся в стадии разработки) параллельных технологий, которые в перспективе предполагается использовать для криптоанализа известных шифров, исследования криптографических хэш-функций, а также при решении задач верификации дискретных автоматов.

Параллельные методы кластерного анализа больших массивов данных (Васильев И.Л., Климентова К.Б., ИДСТУ СО РАН).

Задача кластерного анализа заключается в разбиении исходного множества объектов на заданное количество подмножеств схожих объектов (кластеров) и последующем выборе наиболее “типичного” представителя (медианы) из каждого кластера. Как правило, характеристики объектов задаются числовыми векторами, а за меру схожести принимается расстояние между векторами в определенной метрике. Современные методы кластерного анализа (такие как k -mean, иерархические алгоритмы и др.) способны эффективно обрабатывать массивы данных, состоящие из нескольких тысяч объектов, в то время как в практических задачах требуется кластеризация сотен тысяч и миллионов элементов. Главной проблемой в таких задачах является невозможность вычислять, сохранять и манипулировать всеми расстояниями между объектами.

В Институте динамики систем и теории управления СО РАН разработаны методы кластеризации, основанные на алгоритмах решения известной задачи целочисленного программирования о p -медиане (пи-медиане). Предложен подход, позволяющий эффективно производить кластеризацию нескольких десятков тысяч объектов на современном персональном компьютере в течение 20–30 мин. Однако на практике требуется решение задач гораздо большей размерности и за более короткое время. В связи с этим ведется разработка соответствующего параллельного алгоритма для кластерных систем. Первые тесты такого алгоритма на вычислительном кластере “Blackford” показали, что примеры, содержащие до 100 тысяч объектов, могут решаться в течение нескольких минут. При этом удастся добиться “сверхпараллельного” ускорения работы алгоритма за счет использования всего объема доступной оперативной памяти (160 GB). Несмотря на первые успехи в решении практических задач достаточно большой размерности, в медицине, генетике, экономике зачастую требуется обработка массивов данных еще более высокой размерности, содержащих информацию о миллионах объектов. Для их кластеризации вычислительных мощностей уровня “Blackford” может оказаться недостаточно.

Параллельное решение задач в булевых ограничениях (Опарин Г.А., Богданова В.Г., ИДСТУ СО РАН).

Многие практические задачи могут быть сформулированы как задачи булевой выполнимости (решения системы булевых уравнений). К ним относятся задачи криптографии, логического программирования, доказательства теорем, дедуктивного и абдук-

тивного вывода, нейросетевых вычислений, машинного зрения, планирования действий роботов, планирования вычислений в системах синтеза программ, оптимального проектирования, разработки систем баз знаний, компьютерной графики, моделирования цепей, разработки СБИС, безопасности, составления расписаний, интеллектуального управления, различные задачи на графах, шахматной доске и мн. др. Хорошо известно, что, поскольку булевы системы с точки зрения человеческого восприятия имеют весьма сложную специфическую природу, важной и весьма нетривиальной задачей является построение средств, позволяющих моделировать задачи на языке булевых ограничений. С другой стороны, задачи в булевых ограничениях обладают тяжелыми комбинаторными характеристиками с высокой оценкой сложности. Поэтому необходимы как поиск и реализация методов, наиболее эффективно действующих на отдельных классах булевых систем, интересных с практической точки зрения, так и повышение эффективности решения систем булевых уравнений на основе использования многопроцессорной вычислительной техники. В частности, метод расщеплений позволяет осуществить крупноблочное распараллеливание системы булевых уравнений на любое количество процессоров (ядер), что обеспечивает реальную возможность решения многих сложных теоретических и практических задач на вычислительных кластерах сверхвысокой производительности и вычислительных GRID.

3. Интеллектуальные технологии и инструментальные средства автоматизации крупноблочного параллельного и распределенного программирования

Известно, что проблема автоматизации составления программ является одной из ключевых в параллельном и распределенном программировании. В Институте динамики систем и теории управления СО РАН разработаны высокоуровневые инструментальные средства, позволяющие в проблемноориентированных терминах как конструировать крупноблочные параллельные и распределенные вычислительные процессы, так и автоматически их синтезировать на основе непроцедурных постановок задач на вычислительной модели предметной области. В число этих средств входят следующие.

1. Среда непроцедурного программирования СиКрус [8] для синтеза параллельных программ на языке Fortran-DVM с учетом ресурсных ограничений используемой вычислительной системы и времени исполнения предметных модулей. Выбор языка Fortran-DVM в качестве выходного языка генератора управляющей программы транслятора-синтезатора позволяет обеспечить простоту генерации результирующего кода и удобство его использования прикладным специалистом в плане читаемости, отладки и ручной модификации (в случае необходимости).

2. Инструментальный комплекс ORLANDO Tools [9] для конструирования асинхронных параллельных программ с использованием коммуникационной библиотеки PVM. Комплекс включает совокупность моделей, алгоритмов и программных средств, реализующих в создаваемых программах как параллелизм по управлению, свойственный системам крупноблочного программирования, так и параллелизм по данным, характерный для задач многовариантного анализа и обработки больших массивов данных, допускающих естественную декомпозицию на независимые сегменты.

3. Инструментальный комплекс DISCOMP [14] для организации распределенных пакетов прикладных программ в разнородных вычислительных средах. Функциональ-

ное наполнение таких пакетов может включать модули, реализованные на различных языках программирования, нетиражируемые программные комплексы, а также унаследованное ПО. Комплекс обеспечивает как возможность создания автономных распределенных пакетов прикладных программ, так и интеграцию данного инструментария с другими программными комплексами для автоматизации распределенных вычислений при решении исследовательских задач.

Перечисленные инструментальные средства используются в институтах Иркутского научного центра СО РАН при создании пакетов прикладных программ для решения задач различных предметных областей, например, исследования биоресурсов оз. Байкал, решения систем булевых уравнений и задач оптимального управления, моделирования складского логистического комплекса и др.

4. Средства автоматизации создания и применения кластерных GRID

Традиционные GRID [29–31], организованные на базе кластеров, в большинстве случаев использования обеспечивают возможность удаленного доступа к ресурсам (узлам) вычислительной сети, позволяют определять вычислительные возможности конкретного узла (число процессоров, объем оперативной памяти и т. п.), степень его работоспособности и выполнять на этом узле некоторое независимое задание или обрабатывать один из блоков данных при решении “большой” задачи, допускающей распараллеливание по данным [32]. Однако существуют типы научно-исследовательских задач, для решения которых необходимы дополнительные возможности, в частности:

— предоставление вычислительных услуг нетиражируемых программных комплексов, размещенных в узлах GRID;

— выполнение ряда взаимозависимых заданий, составляющих процесс решения одной общей задачи и требующих интеграции распределенных вычислительных ресурсов на основе автоматического планирования последовательности их использования.

Кроме того, анализ систем управления прохождением задач (СУПЗ) для кластеров показывает, что, как правило, в этих системах реализовано централизованное управление автономными заданиями. Вместе с тем решение задач с большим числом подзадач и информационно-логических связей между ними в значительной степени увеличивает нагрузку на управляющий компьютер кластера и может привести к снижению работоспособности всей вычислительной системы в целом [33].

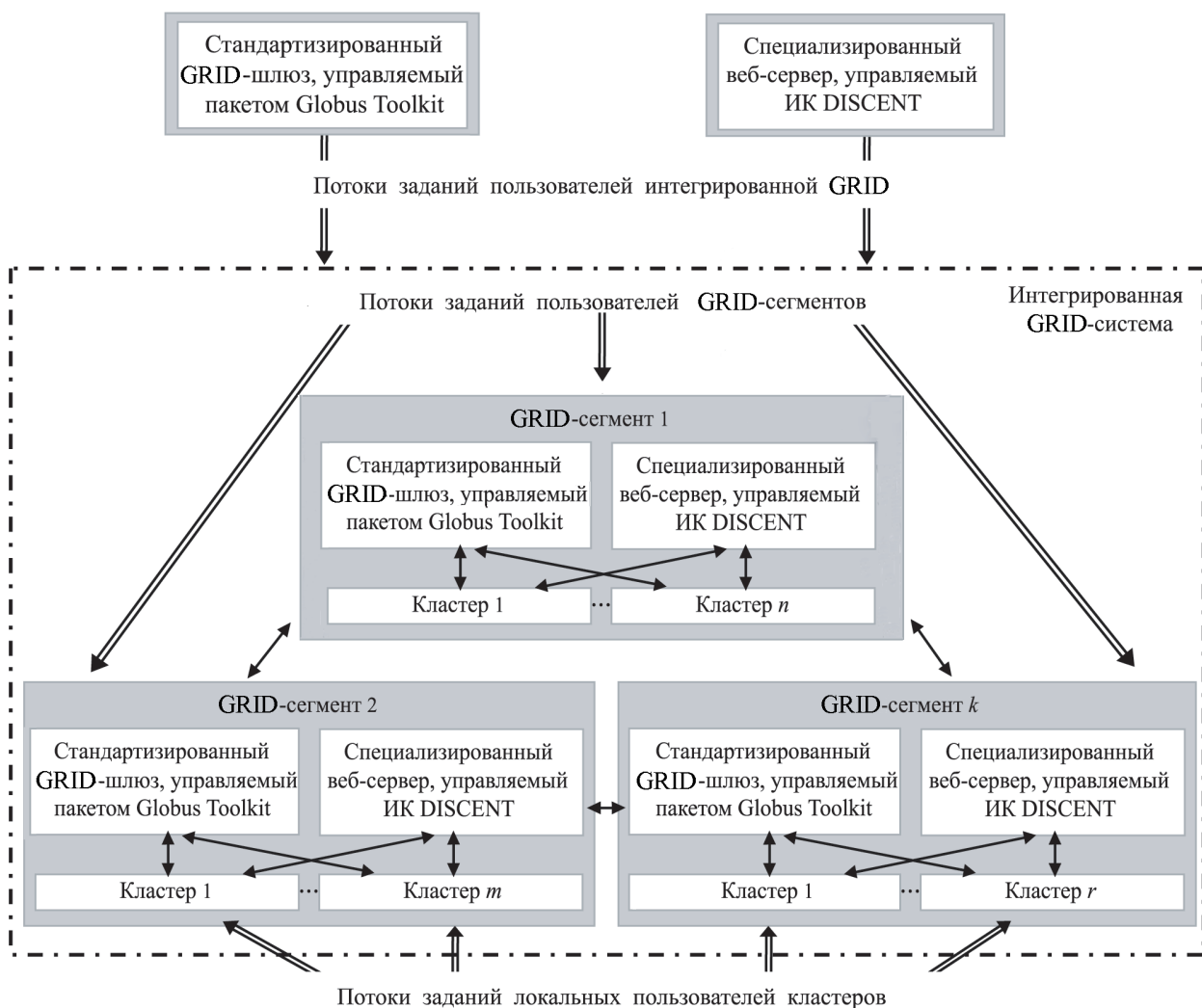
Следует также отметить, что формы ведения вычислительных работ в GRID, обусловленные применением сложного системного ПО, ориентированы главным образом на высококвалифицированных специалистов в области системного программирования. Эти обстоятельства актуализируют необходимость создания таких средств доступа к ресурсам GRID и способов управления распределенными вычислительными процессами [34], которые обеспечивают широкое использование GRID прикладными специалистами. В наших исследованиях предложен новый способ организации кластерных GRID, предоставляющих широкий набор средств решения ресурсоемких научно-исследовательских задач различных типов и реализующих возможность интеграции с другими GRID-системами. В этом направлении в ИДСТУ разработаны следующие средства.

1. Графическая инструментальная среда СИРИУС для анализа эффективности функционирования проектируемых кластерных систем модульного программирования,

обеспечивающая (в отличие от известных) выявление комплексного влияния характеристик как прикладного программного обеспечения, так и аппаратных средств на вероятностно-временные показатели решения прикладной задачи в исследуемой системе. Основой среды являются методы [7] распределенного имитационного моделирования кластерных систем модульного программирования с применением средств динамического анализа программ.

2. Технология и инструментальный комплекс DISCENT [15] для организации GRID с целью решения разнотипных ресурсоемких научно-исследовательских задач. В состав комплекса включены средства, обеспечивающие:

- организацию GRID с использованием стандартизированного базового ПО: пакета Globus Toolkit и согласованных с ним по управлению и форматам данных СУПЗ (PBS, SGE, Condor и др.);
- разработку программных надстроек для решения сложных задач, требующих дополнительных (не реализованных в базовом ПО GRID) возможностей организации вычислительных процессов и децентрализованного управления ими;
- эффективное управление потоками разнотипных заданий в GRID на основе интеллектуального планирования вычислений;



Многоуровневая архитектура интегрированной GRID-системы

— прохождение задач, поступающих с помощью различных средств доступа к ресурсам GRID (включая задачи локальных пользователей узлов GRID) исключительно через кластерные СУПЗ, что позволяет децентрализованным службам управления поддерживать целостность информационной картины, складывающейся в процессе вычислений;

— крупноблочный синтез параллельных распределенных программ для решения информационно-зависимых задач в GRID;

— предварительное проведение математического моделирования вычислительной инфраструктуры GRID с целью оценки эффективности ее функционирования;

— интеграцию нескольких GRID в единую систему с многоуровневой (по управлению потоками заданий) архитектурой путем взаимной сертификации их вычислительных ресурсов средствами пакета Globus Toolkit через стандартизированные GRID-шлюзы (см. рисунок);

— унифицированный прозрачный доступ пользователей (специалистов-“предметников”, решающих свои задачи в GRID) ко всем ресурсам интегрированной GRID-системы через специализированные веб-серверы.

Данный инструментальный комплекс был использован для создания экспериментального кластерного GRID-сегмента ИДСТУ и решения с его помощью ряда практических задач [12]. Результаты вычислительных экспериментов показали устойчивую эффективность функционирования ресурсов GRID под управлением ИК DISCENT (в том числе при интеграции с другими GRID).

Успешное применение разработанных в институте высокоуровневых инструментальных средств позволяет надеяться, что они будут востребованы при организации GRID Сибирского отделения РАН как единой информационно-вычислительной инфраструктуры для поддержки проведения междисциплинарных фундаментальных и прикладных исследований.

5. Подготовка специалистов по параллельным и распределенным вычислениям

Одно из важнейших направлений деятельности ИДСТУ СО РАН — подготовка специалистов в области параллельных и распределенных вычислений. В течение 2005–2009 гг. по данной тематике в рамках специальностей 05.13.01, 05.13.11, 05.13.18 защищены шесть кандидатских диссертаций [35–40]. Для студентов Института математики, экономики и информатики Иркутского государственного университета на базе ИДСТУ открыта специализация “Параллельные и распределенные вычислительные системы”. Ведущими специалистами СКЦ читаются курсы “Параллельные вычислительные системы”, “Программирование систем с распределенной памятью”, “Программирование систем с общей памятью”, “Администрирование кластерных систем”, “GRID-технологии”. Практические занятия проводятся с использованием вычислительного кластера МВС-1000/16 и экспериментального кластерного GRID-сегмента.

В целом в ИДСТУ СО РАН комплексно решается широкий спектр задач, к числу которых относятся создание, эксплуатация, модернизация и развитие базовых программно-аппаратных средств организации высокопроизводительных вычислений; разработка ориентированных на специалистов-“предметников” технологий и инструментальных средств

крупноблочного конструирования и синтеза параллельных программ; создание удобных и развитых средств автоматизации разработки кластерных GRID с учетом их последующего комплексирования в многоуровневые структуры; подготовка высококвалифицированных специалистов, способных решать перечисленные выше задачи.

Список литературы

- [1] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., НОВОПАШИН А.П. Модель поведения интеллектуальных почтовых клиентов в распределенной вычислительной САТУРН-среде // Вычисл. технологии. Т. 7. Вестник КазНУ. № 4 (32). Совместный выпуск. 2002. Ч. 4. С. 21–28.
- [2] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., НОВОПАШИН А.П. Вопросы организации взаимодействия вычислительных модулей в распределенной интеллектуальной САТУРН-среде // Вычисл. технологии. Т. 8. Вестник КазНУ. № 3. Совместный выпуск. 2003. Ч. 2. С. 307–315.
- [3] ОПАРИН Г.А., НОВОПАШИН А.П. Булево моделирование планирования действий в распределенных вычислительных системах // Изв. РАН. Теория и системы управления. 2004. № 5. С. 105–108.
- [4] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., НОВОПАШИН А.П. Оптимальное планирование и децентрализованное управление исполнением схем решения задач в распределенной вычислительной среде // Вычисл. технологии. Т. 9. Вестник КазНУ. № 3 (42). Совместный выпуск. 2004. Ч. 3. С. 246–253.
- [5] ОПАРИН Г.А., НОВОПАШИН А.П. Инструментальная среда синтеза модульных параллельных программ для вычислительных кластеров // Вестник ТГУ. Приложение. 2006. № 17. С. 247–250.
- [6] ОПАРИН Г.А., НОВОПАШИН А.П. Методы и инструментальные средства крупноблочного синтеза параллельных программ на языке Fortran-DVM // Параллельные вычисления и задачи управления: Избр. докл. III Междунар. конф. памяти И.В. Прангишвили. М.: Изд-во ИПУ РАН, 2006. С. 59–63.
- [7] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., АЛЕКСАНДРОВ А.А. Графическая инструментальная среда для описания модели распределенной вычислительной системы // Вестник ИрГТУ. 2006. Т. 3, № 2 (26). С. 35–40.
- [8] ОПАРИН Г.А., НОВОПАШИН А.П. Технология синтеза модульных параллельных программ для DVM-системы // Интеллектуальные системы: Тр. VII Междунар. симп. / Под ред. К.А. Пупкова. М.: РУСАКИ, 2006. С. 468–471.
- [9] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., НОВОПАШИН А.П., ГОРСКИЙ С.А. Инструментальный комплекс ORLANDO TOOLS // Программные продукты и системы. 2007. № 4. С. 63–65.
- [10] ОПАРИН Г.А., НОВОПАШИН А.П. Булевы модели и методы планирования параллельных абстрактных программ // Автоматика и телемеханика. 2008. № 8. С. 166–175.
- [11] ОПАРИН Г.А., НОВОПАШИН А.П. Булевы модели синтеза параллельных планов решения вычислительных задач // Вестник НГУ. Информац. технологии. 2008. Т. 6, вып. 1. С. 53–59.
- [12] ФЕОКТИСТОВ А.Г., КОРСУКОВ А.С. Разработка GRID-системы с децентрализованным управлением потоками заданий // Там же. 2008. Т. 6, вып. 3. С. 147–154.
- [13] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г. Модели и инструментальные средства организации распределенных вычислений // Параллельные вычисления и задачи управления: Материалы IV Междунар. науч. конф. М.: Изд-во ИПУ РАН, 2008. С. 1126–1135.

- [14] ОПАРИН Г.А., ФЕОКТИСТОВ А.Г., СИДОРОВ И.А. Технология организации распределенных вычислений в инструментальном комплексе DISCOMP // Современные технологии. Системный анализ. Моделирование. 2009. № 2. С. 175–180.
- [15] БЫЧКОВ И.В., КОРСУКОВ А.С., ОПАРИН Г.А., ФЕОКТИСТОВ А.Г. Технология организации гетерогенных распределенных вычислительных сред // Научный сервис в сети Интернет: Масштабируемость, параллельность, эффективность: Тр. Всероссийской суперкомпьютерной конф. М.: Изд-во МГУ, 2009. С. 353–359.
- [16] БЫЧКОВ И.В., МАДЖАРА Т.И., НОВОПАШИН А.П. и др. Анализ и перспективы развития инфраструктуры интегрированной информационно-вычислительной сети Иркутского научно-образовательного центра // Вестник НГУ. Информац. технологии. 2008. Т. 6, вып. 1. С. 25–36.
- [17] НОВОПАШИН А.П., СИДОРОВ И.А. Иркутский суперкомпьютерный центр коллективного пользования // Распределенные и кластерные вычисления: Избр. материалы V школы-семинара / Под ред. В.В. Шайдурова. Красноярск: Изд-во ИВМ СО РАН, 2005. С. 94–96.
- [18] ГАЙКАЛОВ И.В., ИЛЬИНА О.В., КИРИЛЬЧИК С.В., СУХАНОВА Л.В. Описание трех микросателлитных локусов байкальского омуля *Coregonus migratorius* (Georgi) // Генетика. 2008. Т. 44, № 3. С. 423–426.
- [19] АННЕНКОВА Н.В., БЕЛИКОВ С.И., БЕЛЫХ О.И. Обнаружение динофлагеллят рода *Gyrodinium* в фитопланктоне и губках озера Байкал путем анализа нуклеотидных последовательностей // Водоросли: Проблемы таксономии, экологии и использование в мониторинге: Материалы II Всероссийской науч.-практ. конф. Сыктывкар: Изд-во Института биологии Коми НЦ УрО РАН, 2009. С. 324–327.
- [20] BLASCHKE D., BUBALLA M., RADZHAPOV A.E., VOLKOV M.K. Nonlocal quark model beyond mean field and QCD phase transition // Nucl. Phys. B. Proc. Suppl. 2010. Vol. 198, Iss. 1. P. 51–54.
- [21] MYASNIKOVA A.S., RADZHAPOV E.A., MYSOVSKY A.S., SHAGUN V.A. Impurity Luminescence in BaF₂ Crystals // IEEE Trans. Nuclear Sci. 2008. Vol. 55, Iss. 3. P. 1133–1137.
- [22] MYSOVSKY A., RADZHAPOV E. The refinement of self-trapped excitons structure in CaF₂ and SrF₂ crystals: An ab initio study // Book of Abstr. of SCINT-2009. Jeju, Korea, 2009. P. 43.
- [23] ЛАЗЕБНЫХ В.Ю., МЫСОВСКИЙ А.С., СИНИЦА Л.Н. Теоретическое исследование эффекта спин-селективной адсорбции молекул воды на поверхности MgO // Оптика и спектроскопия. 2009. Т. 107, № 4. С. 606–611.
- [24] ЗАЙКИН О.С., СЕМЕНОВ А.А. Технология крупноблочного параллелизма в SAT-задачах // Проблемы управления. 2008. № 1. С. 43–50.
- [25] СЕМЕНОВ А.А., ЗАЙКИН О.С. Неполные алгоритмы в крупноблочном параллелизме комбинаторных задач // Вычисл. методы и программирование. 2008. Т. 9, № 1. С. 112–122.
- [26] СЕМЕНОВ А.А., ЗАЙКИН О.С., БЕСПАЛОВ Д.В., УШАКОВ А.А. SAT-подход в криптоанализе некоторых систем поточного шифрования // Вычисл. технологии. 2008. Т. 13, № 6. С. 134–150.
- [27] ОПАРИН Г.А., БОГДАНОВА В.Г. РЕБУС — интеллектуальный решатель комбинаторных задач в булевых ограничениях // Вестник НГУ. Информац. технологии. 2008. Т. 1, вып. 6. С. 60–68.
- [28] ОПАРИН Г.А., БОГДАНОВА В.Г., МАКЕЕВА Н.Г. Инструментальная среда параллельного решения систем булевых уравнений // Современные технологии. Системный анализ. Моделирование. 2009. № 3. С. 62–68.

- [29] Коваленко В.Н., Корягин Д.А. Организация ресурсов ГРИД. М., 2004 (Препр. РАН ИПМ им. Келдыша. № 63).
- [30] Демичев А., Ильин В., Крюков А. Введение в ГРИД-технологии. М., 2007 (Препр. РАН НИИЯФ МГУ. № 11/832).
- [31] Коваленко В.Н. Комплексное программное обеспечение ГРИДа вычислительного типа. М., 2007 (Препр. РАН ИПМ им. Келдыша. № 10).
- [32] Воеводин В.В. Решение больших задач в распределенных вычислительных средах // Автоматика и телемеханика. 2007. № 5. С. 32–45.
- [33] Каляев И.А., Гайдук А.Р., Капустян С.Г. Распределенные системы планирования действий коллективов роботов. М.: Янус-К, 2002. 292 с.
- [34] DURFEE E.H. Distributed problem solving and planning // Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence / Ed. G. Weiss. Cambridge: MIT Press, 1999. P. 121–164.
- [35] Новопашин А.П. Методы и инструментальные средства крупноблочного синтеза параллельных программ для вычислительных кластеров: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2005. 26 с.
- [36] Горский С.А. Инструментальный комплекс для организации параллельных вычислений в интеллектуальных пакетах прикладных программ: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2008. 20 с.
- [37] Корсуков А.С. Инструментальный комплекс для разработки и применения гетерогенных распределенных вычислительных сред: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2009. 18 с.
- [38] Заикин О.С. Параллельная технология решения SAT-задач и ее реализация в виде пакета прикладных программ: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2009. 19 с.
- [39] Александров А.А. Инструментальный комплекс распределенного имитационного моделирования кластерных систем модульного программирования: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2009. 18 с.
- [40] Сидоров И.А. Технология и инструментальные средства организации распределенных пакетов прикладных программ в разнородных вычислительных средах: Автореф. дис. ... канд. техн. наук. Иркутск, ИДСТУ СО РАН, 2009. 18 с.

Поступила в редакцию 22 марта 2010 г.