

Химический GRID в ИПХФ РАН: 5 лет развития

В.М. Волохов¹, Д.А. Варламов^{1,2}, А.В. Пивушков¹, Г.А. Покатович¹, Н.Ф. Сурков¹

¹Институт проблем химической физики РАН, Московская обл., Черноголовка

²Институт экспериментальной минералогии РАН, Московская обл., Черноголовка

1. Введение

Современное мировое состояние вычислительной химии характеризуется использованием сверхмощных параллельных и распределенных вычислительных ресурсов для решения задач различных классов. Вычислительная и квантовая химия являются одними из наиболее заинтересованных в GRID вычислениях отраслями науки.

Для проведения крупномасштабных вычислений в области вычислительной и квантовой химии и сопряженных областей (газодинамики экстремальных состояний, моделирования сложных биологических систем, строения вещества, нанотехнологий, разработки новых лекарственных препаратов и т.п.) требуется проведение высокоинтенсивных параллельных и распределенных расчетов. Например, некоторые задачи оптимизации молекулярных структур требуют выполнения до 10^9 отдельных расчетов. Для подобных задач необходимо развитие и применение технологий GRID в области вычислительной и квантовой химии для организации распределенных вычислений.

Крупномасштабные квантово-химические расчеты – одно из основных научных направлений ИПХФ РАН. Эти расчеты выполняются с использованием авторских программ, "open source" пакетов (CPMD, Dalton-2, GAMESS-US, NAMD, ABINIT и др.), а также лицензионных программ (Gaussian-98,-03, Морас2002, MolPro). Институт располагает богатейшей в России библиотекой параллельных квантово-химических и молекулярно-динамических программ. Работы с системами распределенных вычислений в ИПХФ РАН были начаты в 2004 году в рамках программы № 21 фундаментальных исследований Президиума РАН «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID» и продолжены в рамках Федеральной целевой научно-технической программы "Исследования и разработки по приоритетным направлениям развития науки и техники" по проекту "Создание комплекса пакетов прикладных программ для моделирования сложных научных и промышленных задач на суперкомпьютерных системах терафлопного уровня и в распределенных вычислительных средах", а также в рамках программы Союзного Государства «СКИФ-GRID».

Основными задачами авторов в эти годы стало развитие двух основных направлений: 1) адаптацию наиболее востребованного прикладного ПО в области вычислительной (прежде всего квантовой) химии к работе в GRID инфраструктуре и обеспечение широкого доступа пользователей к работе с ним с использованием самых различных методов и технологий; 2) развитие ресурсного узла GRID, выступающего как в роли полигона для проведения вычислительных экспериментов в данной области, так и в роли средства для решения реальных фундаментальных и научно-практических задач.

2. Основные типы и классы задач вычислительной химии

Квантово-химические расчеты являются важнейшим звеном при проведении исследований в области строения вещества, наноматериалов, физики твердого тела, биофизики и всех научных дисциплин, связанных с исследованием электронной структуры вещества и его строения.

Многолетний опыт позволил разделить квантово-химические задачи на два основных вычислительных типа: 1) распадающиеся на совокупность практически независимых заданий; 2) задачи, представляющие собой единый вычислительный процесс.

Характерная особенность задач первого типа состоит в том, что, распределяя независимые задания на множество узлов или небольших кластеров (каждый кластер – 10-100 процессоров, задание может исполняться как параллельное), можно добиться высокой эффективности использования вычислительных ресурсов. При этом возможно использование больших вычислительных полигонов (до 10^3 процессоров) как в едином кластере, так и в совокупности удаленных кластеров.

Задачи второго типа представляют собой существенную проблему, т.к. эффективность их решения непосредственно связана с эффективностью распараллеливания вычислительного процесса и высокими требованиями к ресурсам узла, особенно по памяти. Для программы Gaussian известно эмпирическое правило: масштабируемость пропорциональна кубическому корню из числа процессоров. Для программы GAMESS масштабируемость существенно лучше и для нескольких десятков процессоров остается практически линейной. Для характерных задач исследования наноструктур и молекулярных кристаллов необходимо несколько тысяч процессоров и процессорное время порядка месяца, причем желательно использование кластеров терафлопного уровня.

3. Основные направления работ по применению GRID технологий в химии

Нами проводилась экспериментальная проверка и апробация возможности использования GRID ресурсов для реальных расчетов на стандартных пакетах прикладных программ (в том числе и параллельных), используемых в вычислительной химии, таких, как GAMESS, Gaussian, Dalton2, CPMD и других, а также различных авторских программ, разработанных в ИПХФ и ИЦЧ РАН.

Для адаптации в распределенных вычислительных средах были выбраны следующие прикладные программные пакеты:

1. **GAMESS-US** (<http://www.msg.ameslab.gov/GAMESS/>);
2. **Dalton-2** (<http://www.kjemi.uio.no/software/dalton/dalton.htm>);
3. **CPMD** (<http://www.cpmc.org/>);
4. **NAMD** (<http://www.ks.uiuc.edu/Research/namd/>);
5. **Gaussian03** (<http://www.gaussian.com/>);
6. **Авторские программы**, включающие многопараметрические задачи из области квантовой химии и молекулярной динамики, параллельные газодинамические программы моделирования (на молекулярном уровне) процесса образования ударных волн и др.

Для всего выбранного ПО был проведен детальный анализ модульной структуры квантово-химического кода и изучены особенности работы различных реализаций однопроцессорных и параллельных версий, определены стратегии реализации выбранных типов квантово-химических вычислений применительно к распределенным средам.

3.1 Начальные варианты расчетов в распределенных средах

В 2004-2005 годах для построения первичного полигона ИПХФ по применению параллельных и распределенных вычислений нами была выбран ряд модельных задач из области квантовой химии, например, процесс туннельного прохождения протона через потенциальный барьер, параметры которого периодически зависят от времени, а параметрами являются частота и амплитуда излучения. Задачи имеют высокую вычислительную сложность, однако вычисления в каждой точке сетки происходят независимо друг от друга, что позволяет разбить область вычислений на множество непересекающихся подобластей и на каждой из них запускать задачи на различных процессорах. В качестве вычислительных систем были выбраны распределенные комплексы Condor и X-Com. На языке Perl были написаны процедуры для “нарезки” сетки для вычислений и слияния насчитанных результатов. Решалась задача с входными данными, описывающими потенциальную кривую и параметры периодических колебаний внеш-

него поля. Область решения задачи была разбита на 30 областей по ~4200 точек. Для одиночного ПК оцениваемое время решения составило бы около 5-6 тысяч часов.

Система Condor (<http://www.cs.wisc.edu/condor>) позволяет использовать чрезвычайно разнородные вычислительные ресурсы (преимущественно внутри локальных сетей) для решения одной параллельной задачи и организовывать многократное выполнение задания на массивах разных входных данных. Для запуска «пучка» заданий в системе Condor для вычислений на «нарезанной» сетке использовалось свойство системы Condor организовать многократное выполнение задания на разных входных данных – для этого нужно специальным образом оформить задание, так, чтобы для каждого запуска было предусмотрено свое множество входных и выходных файлов. Запуск «пучка» заданий осуществлялся стандартными средствами Condor'a. Condor полностью автоматизирует процесс распределения заданий. Одновременно система обеспечивает ведение дисциплины нескольких уровней приоритетов рассмотрения заявок: приоритет назначения ресурсов, приоритет использования и приоритет среди машин, удовлетворяющих одинаковым заявкам.

Вторая исследованная система – система метакомпьютинга X-Com (разработка НИВЦ МГУ, <http://x-com.parallel.ru>) позволяет в распределенном режиме решать задачи с высокой вычислительной сложностью также на гетерогенных платформах. Для проведения вычислительных экспериментов на системе X-Com использовалась тот же тип задач, что и в варианте с Condor. Для адаптации задачи к системе X-Com были написаны (в содружестве с НИВЦ МГУ) клиентские и серверные ее части. Серверная часть осуществляла разбиение области вычислений на подобласти и передачу их клиентам. Клиентская часть на каждом вычислительном узле запускала задачу на счет и по завершении отправляла серверу файлы результатов. Для удобства работы был написан простой WWW-интерфейс, позволяющий вводить числовые параметры задачи в WWW-форме, а также утилита командной строки, принимающая на вход файл со значениями параметров в стандартном для задачи формате.

Проведенные вычислительные эксперименты с использованием до 40 процессоров для Condor и до 400 – для X-Com показали, что для таких задач квантовой химии функциональные возможности обеих систем достаточны и примерно равны, а развертывание систем не представляет большой сложности.

Были созданы методы запуска «пучков» независимых заданий для использования всех доступных распределенных ресурсов для широкого класса многопараметрических задач вычислительной химии. При этом полная задача разбивается на огромное количество независимых подзадач (каждая определяется группой значений совокупности параметров, обычно до 10^4 , в перспективе – до 10^7 «атомарных» заданий на задачу). Была разработана методика расчета задач и получения результатов методом запуска «пучков» заданий на всех доступных ресурсах с использованием низкоуровневых интерфейсов применяемого middleware. Для выбранных областей данных авторскими скриптами производится «нарезка» областей данных, формирование пула независимых заданий, создание очередей запуска и отправки заданий на брокер ресурсов. После запуска периодически запускаемые (средствами ОС) скрипты ведут мониторинг выполнения заданий, контроль таймаутов, перезапуск неудачных заданий и сбор результатов выполненных заданий (с использованием базы данных и таблиц в ней, контролирующей состояние заданий – «ожидание», «запуск», «выполнение» и т.д.). По окончании расчетов проводится сборка «атомарных» результатов в единый выходной файл. Для части задач (требующих значительного числа параллельных независимых расчетов) дополнительно созданы авторские механизмы по разбиению областей данных (или расчетов) на большие независимые подсетки или независимые задания, передачи всех их низкоуровневым интерфейсам gLite или Unicore с последующим запуском на парал-

лельных узлах и «сборки» финальных результатов.

3.2 Распределенные среды, используемые на текущем этапе

Однако примененные технологии были слабо увязаны с основным направлением развития инфраструктуры GRID, поэтому был сделан переход на разработанную в CERN среду gLite (<http://glite.web.cern.ch/gLite>, на ранних стадиях LCG-2), на базе которой и был создан ресурсный узел консорциума EGEE-RDIG (<http://www.egee-rdig.ru>). В 2008 году в рамках проекта был создан ресурсный сайт полигона СКИФ-ГРИД (<http://skif-grid.botik.ru>) для работы в среде Unicore (<http://www.unicore.eu>).

Для выбранных прикладных пакетов созданы и протестированы на реальных задачах низкоуровневые интерфейсы для запуска их в распределенных вычислительных средах gLite и Unicore. Данные интерфейсы включают набор скриптов по формированию исходящих заданий, запуску через брокер ресурсов на удаленных ресурсах, мониторингу выполнения задач, возвращению полученных результатов с удаленных ресурсов и «сборку» окончательных результатов на интерфейсе пользователя. Реализованы интерфейсы для однопроцессорных и параллельных (SMP, сокетные, MPI) вариантов указанного ПО. На ресурсных GRID узлах ИПХФ, использованных в качестве удаленного распределенного ресурса, проведены запуски указанного прикладного ПО в рамках инфраструктур BO RGSTEST RDIG (с прохождением задач через центральный брокер ресурсов RDIG) и СКИФ-Полигона. Запуски всего адаптированного ПО проводились в разных режимах и конфигурациях (с разным количеством востребованных процессоров и использованием разных вариантов параллельных расчетов). Были изучены варианты совмещения различных вариантов распараллеливания (например, SMP+MPI) вычислений применительно к некоторым прикладным пакетам (пакеты Dalton-2 и CPMD). После ряда вычислительных экспериментов была проведена коррекция созданных низкоуровневых интерфейсов и окончательная их оптимизация. Были скорректированы проблемы запуска и работы параллельных (SMP, сокетные, MPI-1,2) вариантов указанного ПО на различных типах ресурсных узлов (разные системы PBS и параллельные среды).

Следует отметить, что большинство указанных прикладных пакетов вычислительной химии отличаются сложностью конфигураций и повышенными требованиями к среде выполнения, особенно для проведения параллельных расчетов. Далеко не всегда возможна перенастройка ресурсных узлов распределенных сред под нужды подобных пакетов или предустановка их на ресурсных узлах. Поэтому авторами был разработан метод создания виртуальных «контейнеров», перемещаемых стандартными средствами распределенного middleware и запускаемых на удаленном ресурсном узле распределенной среды. В рамках вычислительной химии подобная концепция становится весьма востребованной, поскольку крупные пакеты (типа GAMESS, Gaussian, Dalton, CPMD, NAMD) требуют значительной настройки ОС всех участвующих в вычислениях узлов. Применение таких «контейнеров» позволяет передавать заранее настроенную среду как единое задание, не требующее дополнительного конфигурирования и сложной процедуры установки и настройки, производимых, как правило, вручную администратором кластеров. На удаленный ресурсный узел сети GRID через брокер ресурсов передается главный скрипт и упакованный модуль, содержащий исполняемые файлы приложения, необходимые MPI-2 библиотеки и скрипты запуска и выполнения приложения, т.е. динамически сформированный «виртуальный контейнер». После доставки «контейнера» на ресурсный узел проводится ряд шагов по месту исполнения, позволяющие произвести установку необходимого ПО и библиотек, перенастройку окружения и запуск «кольца» серверов mpd на узле GRID с последующим запуском параллельного приложения как обычного распределенного задания с финальной передачей результатов на брокер ресурсов и затем пользователю и «очисткой» ресурсного узла.

Так могут быть решены проблемы установки, настройки, несовместимости с операционной системой и другими программами.

3.3 WWW портал Grid Enabled Chemical Physics (GECР) ИПХФ РАН

Упрощенный доступ к ресурсам и сервисам является одним из наиболее важных компонентов GRID системы, так как является ключевым связующим звеном между GRID средой и конечным пользователем. Поэтому необходимо создание GRID порталов, объединяющих GRID и Web сервисы. Это среда, которая позволяет пользователям через web-браузер получить доступ к обширным GRID ресурсам и сервисам.

Портальный интерфейс также очень важен потому, что с помощью него пользователь даже начальной подготовки и уровня знаний может без проблем начать работу в GRID. Архитектура GRID портала основана на идее, что портальная система является контейнером для пользовательских интерфейсов (инструментов, клиентов), обеспечивающих работу с GRID службами. Преимущество данной архитектуры в том, что она является очень гибкой и удобной в разработке и развитии порталов, позволяет встраивать в портал интерфейсы новых GRID служб и изменять существующие. Портальные сервисы контролируют и визуализируют пользовательский интерфейс. Все информационные материалы, представляемые пользователю, поступают от web сервисов из разных источников.

Портал Grid Enabled Chemical Physics (GECР, <http://grid.icp.ac.ru>) объединяет несколько Web-интерфейсов:

1. Квантово-химический комплекс GAMESS для теоретического исследования свойств химических систем *ab initio*;
2. Вычисление многопараметрических функций, под которой следует понимать целый класс нераспределенных задач химической физики, обладающих свойством параллелизма по данным (Data Parallel).

Данные интерфейсы позволяют определять входные параметры и условия (включая загрузку данных и конфигурационных файлов), формировать сложные первичные файлы запуска, производить (при условии сертификации пользователя) запуск данного ПО в распределенных средах, осуществлять мониторинг выполнения заданий и сбор результатов. Интегрирована также технология работы через web-интерфейс с «пучками» независимых заданий на «нарезаемых» областях данных.

4. Заключение

Наши работы позволили создать в рамках технологий GRID вычислительную среду для проведения крупномасштабных расчетов в области вычислительной химии. Это позволило достигнуть нового уровня расчетов в области вычислительной химии:

- создан комплекс адаптированных к среде GRID прикладных программных пакетов вычислительной химии с интерфейсами различного уровня (от низкоуровневых интерфейсов вплоть до Web-портала),
- разработаны новые методики вычислений (включая применение широкого спектра средств виртуализации) в распределенных и параллельных средах применительно к прикладному ПО вычислительной химии;
- создан ресурсный центр (включающий ресурсные узлы на базе разного middleware) для проведения вычислительных экспериментов в этой предметной области, объединяющий как вычислительные узлы для входящих задач в средах gLite и Unicore, так и пользовательские интерфейсы к различным распределенным средам.