

<sup>1</sup>В.М. Волохов, <sup>1</sup>А.В. Пивушков, <sup>1</sup>А.В. Волохов, <sup>1,2</sup>Д.А. Варламов

<sup>1</sup>Институт проблем химической физики РАН, Московская обл., Черноголовка

<sup>2</sup>Институт экспериментальной минералогии РАН, Московская обл., Черноголовка

## РЕАЛИЗАЦИЯ НЕСКОЛЬКИХ НЕЗАВИСИМЫХ РЕСУРСНЫХ ГРИД-САЙТОВ НА ЕДИНОМ ФИЗИЧЕСКОМ ПРОСТРАНСТВЕ КЛАСТЕРА

При организации работ в различных распределенных сетях периодически возникает ситуация, когда необходимо провести тестирование настроек и функциональности ресурсного грид-сайта в условиях распределенного вычислительного полигона еще ДО формирования и подключения собственно высокопроизводительного вычислительного ресурса (крупного кластера, «фермы», сети рабочих станций и т.п.). Иногда необходимо проведение различных вычислительных экспериментов в условиях распределенных сред, при том, что использование основных вычислительных ресурсов организации в таких случаях не всегда желательно. Кроме того, достаточно часто необходимо провести тестовые запуски исходящих задач (особенно в случае новых версий сложных прикладных пакетов) на собственных ресурсах (используемых как удаленные) для отладки, выявления ошибок и проблем, связанных с выполнением задач в грид-средах. Это не всегда возможно как из-за трудоемкости/невозможности перенастройки действующего ресурса (включая конфликты приложений), так и из-за отсутствия такового (размещенного в рамках выбранной грид-сети). Для преодоления подобных проблем авторами была разработана методика размещения небольших ресурсных грид-сайтов разных вычислительных полигонов на едином физическом пространстве небольшого кластера.

**Постановка задачи.** Ранее в ИПХФ РАН реализованы и успешно эксплуатируются несколько ресурсных сайтов различных распределенных грид-полигонов [1–3]:

1) ресурсный узел консорциума EGEE-RDIG (Enable GRID for E-science и Russian Data Intensive GRID, <http://www.egee-rdig.ru>) на основе распределенной среды gLite (<http://glite.web.cern.ch>), работы ведутся в рамках виртуальной организации (ВО) RGSTEST;

2) ресурсный сайт категории «А» для работы в рамках созданного в России крупномасштабного вычислительного полигона СКИФ-Полигон (<http://skif-grid.botik.ru>) на базе промежуточного ПО Unicore (<http://www.unicore.eu>);

3) сайт Национальной нанотехнологической сети (ГридННС, <http://www.ngrid.ru>), среда Globus Toolkit 4, <http://www.globus.org>).

Созданные ресурсные сайты в первую очередь позволяют решать *входящие* задачи как общего характера, так и с использованием прикладных квантово-химических пакетов.

В состав ресурсных сайтов помимо ряда грид-сервисов входит также комплекс клиентских интерфейсов различных уровней для взаимодействия адаптированного квантово-химического прикладного ПО с указанными грид-средами. Они позволяют запускать *исходящие* задачи вычислительной химии на распределенных ресурсах, обеспечивая возможность формирования заданий, запуск на удаленных сайтах (через брокеры ресурсов или в режиме point-to-point -> P2P), мониторинга прохождения заданий, сбора результатов и статистики.

Работа в рамках ВО RGSTEST сети EGEE-RDIG обеспечивает доступ к вычислительным мощностям до 800 процессоров и дисковым массивам порядка 8–15 терабайт в нескольких географических зонах (Москва, Протвино, Харьков, Черноголовка и др.). Разнородность узлов данной ВО позволяет легко варьировать параметры запускаемых задач, ориентируясь на различные типы ресурсов. Использование подобного полигона обеспечивает проведение достаточно масштабных вычислительных экспериментов как научного, так и прикладного характера.

Ресурсный сайт для среды Unicore позволяет выполнять входящие задачи сертифицированных пользователей СКИФ-Полигона, производить мониторинг задач и передавать полу-

ченные результаты пользователям. Обеспечена возможность мониторинга состояния сайта извне. Клиентский интерфейс обеспечивает запуски исходящих задач через среду Unicorn на собственном ресурсном сайте ИПХФ (в роли удаленного ресурса – <https://unicorgw.icp.ac.ru:8080>) и на доступных ресурсных узлах СКИФ-Полигона.

В рамках создаваемой с 2010 года Национальной нанотехнологической сети (ГридННС) предоставлен доступ к ресурсам с общим числом CPU более 8000 (<http://mon.ngrid.ru/stats?page=usage>) и большим количеством виртуальных организаций, в том числе поддерживающих квантово-химические расчеты (на базе сайта ИПХФ создана и функционирует ВО nanochem).

Исходя из опыта эксплуатации созданных ресурсных сайтов, авторами было принято решение о создании небольшого экспериментального кластера для проведения вычислительных экспериментов и разработку грид-приложений без вовлечения основных ресурсов ИПХФ РАН. Кроме того, данная работа позволила в полной мере оценить степень применимости новейших технологий виртуализации ресурсов для обслуживания грид-сервисов и грид-интерфейсов.

**Реализация вычислительного кластера.** Был установлен и настроен вычислительный кластер из 4 двухядерных расчетных узлов (с перспективой ближайшего роста до 8 узлов) с общим количеством ядер 16 и трех сопровождающих управляющих машин, опирающихся на сеть Gigabit-Ethernet. На расчетных узлах был установлен дистрибутив Scientific Linux 5.4 Wogon, поскольку наличие данного дистрибутива требуется для расчетных узлов ресурсного сайта gLite, а для других распределенных сред выбор ОС (варианты Linux) не принципиален. Для проведения вычислительных экспериментов на расчетные узлы был установлен и настроен ряд прикладных параллельных квантово-химических пакетов (Gamess-US, Gaussian, Dalton-2, NAMD). Для обеспечения параллельных вычислений использовались как сокетные версии пакетов, так MPI-варианты. На управляющих машинах был установлен гипервизор KVM (Kernel-based Virtual Machine, <http://www.linux-kvm.org>). На ста-

дии изучения потенциальных вариантов гипервизоров для управляющих машин были протестированы несколько свободно распространяемых гипервизоров виртуальных машин – VirtualBox, VMware, Xen [4], однако, исходя из простоты администрирования, минимума накладных расходов ресурсов и устойчивости работы под нагрузкой в качестве гипервизора был выбран KVM. После установки гипервизора на управляющих машинах были установлены следующие сервисные виртуальные машины:

1) для Computing Element среды gLite 3.1 (полигон EGEE-RDIG) – Scientific linux 4,5 и соответствующие сервисы типа lcg-se, сервисов авторизации и мониторинга;

2) для среды Unicore 6.2 (сайт СКИФ-Полигона) – ОС Ubuntu 9.10, сервисы – шлюз (Gateway); серверный контейнер (Unicore/X), интерфейс к целевой системе (TSI), авторизационный сервис-пользовательская база данных – XUUDBS, пользовательский интерфейс (UI);

3) для среды Globus Toolkit 4 (полигон ГридННС) – ОС CentOS 5.4, сервисы MDS, GRAM, GridFTP, RFT, User Interface.

Выбор соответствующих ОС для управляющих узлов обусловлен либо требованиями дистрибутивов распределенного ПО (как, например, для gLite), либо рекомендациями разработчиков, либо простотой администрирования. На все управляющие машины были установлены серверные компоненты пакета управления заданиями PBS/Torque (<http://www.clusterresources.com>), тогда как на расчетные узлы была установлена клиентская часть данного пакета. Поскольку все распределенные middleware требуют наличия своих собственных очередей PBS, было принято решение о настройке трех одновременно работающих экземпляров pbs\_tom на расчетных узлах с соответствующим набором очередей заданий. В таком варианте каждая управляющая машина связывается с расчетными узлами по уникальному порту и имеет дело только со своими заданиями. Недостатком данного подхода является невозможность (пока) правильного учета ресурсов, используемых расчетным узлом, однако для экспериментальных работ и отладки прикладного ПО это вполне приемлемо. Преимуществами же являются:

1) необходимость ОДНОКРАТНЫХ установки и настройки ПО для решения входящих задач (особенно прикладных, поскольку часто установка прикладных пакетов оборачивается непредвиденными трудозатратами);

2) простота администрирования расчетных узлов;

3) экономия ресурсов (включая прежде всего электроэнергию);

4) повышенный коэффициент загрузки за счет лучшей утилизации CPU.

Преимуществами использования виртуальных машин в качестве управляющих является повышенная надежность ресурсного узла. В случае поломки физического узла копия виртуальной машины (размещенная на другом узле) может быть запущена в считанные минуты. При наличии достаточного объема ресурсов (особенно оперативной памяти) все три виртуальные машины могут быть размещены на одном физическом узле (подобные опыты ставились).

Отметим, что попутно на управляющих машинах можно разместить дополнительные виртуальные машины, отвечающие за различные нересурсоемкие сервисы сети, поскольку их влияние на основные сервисы незначительно.

В ИПХФ разработана методика одновременного размещения нескольких ресурсных сайтов разных распределенных полигонов на едином физическом пространстве одного кластера. Это осуществляется применением технологий виртуализации ресурсов с использованием гипервизора KVM и разделенных экземпляров пакетов управления заданиями PBS/Torque. Подобная методика позволяет проводить обширные вычислительные эксперименты и разрабатывать сложные грид-приложения для различных грид-полигонов без вовлечения основных вычислительных ресурсов организации. Введение подобных вариантов использования кластеров позволяет участвовать в работе нескольких грид-полигонов, значительно повышает надежность управляющих узлов ресурсных сайтов, снижает трудоемкость администрирования сайтов, понижает расходы на поддержку грид-инфраструктуры.

## Список литературы

1. Технологии ГРИД в вычислительной химии / В.М. Волохов [и др.] // Вычислительные методы и программирование. – 2010. – Т. 11, № 1. – с. 42–49.
2. Вычислительная химия в грид-средах / В.М. Волохов [и др.] // Computational Chemistry in the Grid Environments // Distributed Computing and Grid-technologies in science and education: 4<sup>th</sup> Int. Conf. Dubna, JINR, 2010. P.143–144.
3. GRID и вычислительная химия / В.М. Волохов [и др.] // Вычислительные методы и программирование. – 2009. – Т. 10, № 2. – С.78–88.
4. Виртуальные вычислительные среды: использование на GRID полигонах / В.М. Волохов [и др.] // Вестн. ЮУрГУ. Серия «Математическое моделирование и программирование». – 2009. – № 17 (150). Вып. 3. – С. 24–35.

**Н.И. Гаврилов, А.А. Белокаменская**

Нижегородский государственный университет им. Н.И. Лобачевского

### **ОРГАНИЗАЦИЯ ПОТОКОВЫХ ВЫЧИСЛЕНИЙ НА GPU В ЗАДАЧЕ СТЕРЕОВИЗУАЛИЗАЦИИ ТОМОГРАММ**

В научной визуализации часто приходится иметь дело со скалярными полями данных, заданными в трехмерном пространстве. Данные могут быть получены самыми различными способами – численным экспериментом, сканированием с помощью магнитного резонанса, компьютерной томографией (КТ), сканированием с помощью ультразвука. Например, результатом КТ является множество слоёв, т.е. двумерных массивов данных, которые вместе образуют трехмерное скалярное поле. Объем данных эксперимента значителен. Например, объем одного исследования в медицинской томографии обычно равен 0,25–1 Гб. Обеспечение реального времени визуализации означает вывод на экран порядка 25 и более кадров в секунду.