

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. ЛОМОНОСОВА**

---

**НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ  
ЯДЕРНОЙ ФИЗИКИ имени Д.В. СКОБЕЛЬЦЫНА**

А.В. Бережной, В.Ф. Еднерал, В.А. Ильин,  
А.П. Крюков, Г.Б. Шпиз

**ТЕХНОЛОГИЯ ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ  
РЕСУРСОВ ГРИДНС ПОЛЬЗОВАТЕЛЯМИ СИСТЕМЫ  
СКИФ-ГРИД**

Препринт НИИЯФ МГУ № 2009-10/854

Москва, 2009 г.

УДК

**А.В. Бережной, В.Ф. Еднерал, В.А. Ильин,  
А.П. Крюков, Г.Б. Шпиз**

**e-mail: aber@trtk.ru**

**ТЕХНОЛОГИЯ ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ГРИДННС  
ПОЛЬЗОВАТЕЛЯМИ СИСТЕМЫ СКИФ-ГРИД**

Препринт НИИЯФ МГУ № 2008-10/854

Аннотация

*Рассматривается подход к сопряжению вычислительных ресурсов нанотехнологической сети ГридННС с другими грид-системами на основе специальных транзитных узлов в грид-системе пользователя, предназначенных для передачи заданий пользователя в ГридННС и контроля их выполнения. Таким образом, обеспечивается совместное использование ресурсов ГридННС как пользователями ГридННС, так и пользователями других грид-систем (в частности, системы СКИФ-ГРИД). Настоящая работа является развитием разработанной ранее технологии запуска заданий glite на ресурсах СКИФ-ГРИД.*

**A.V. Berezhnoy, V.F. Edneral, V.A. Ilyin,  
A.P. Kryukov, G.B. Shpiz**

**USAGE OF GRIDNNN CALCULATION RESOURCES  
BY SKIF-GRID SYSTEM USERS**

Preprint MSU SINP N 2009-10/854

Abstract

*The coupling of nanotechnology net calculation resources (GRIDNNN) with other GRID systems are discussed. The coupling is implemented with the help of special transaction nodes of user GRID system, which are intended for transaction and monitoring of user jobs submitted into GRIDNNN. Thus GRIDNNN users and users of other GRID systems are provided for the shared GRIDNNN and other GRID resources (particularly SKIF GRID resources). This work is devoted to further development of the earlier designed method for submitting glite jobs to SKIF-GRID resources.*

© А.В. Бережной, В.Ф. Еднерал, В.А. Ильин, А.П. Крюков, Г.Б. Шпиз

© НИИЯФ МГУ, 2009, <http://www.sinp.msu.ru>

## ***Общие принципы интеграции внешних вычислительных ресурсов в грид-системы.***

Под сопряжением грид-систем с внешними вычислительными ресурсами (в частности ресурсами других грид-систем) понимается обеспечение доступа пользователей этих систем к соответствующим ресурсам как для выполнения заданий, так и для получения необходимой информационной поддержки (получение информации о наличии и состоянии ресурсов, о текущем состоянии задач и т.п.). Такое сопряжение грид-систем является важным этапом на пути их объединения в мощные вычислительные структуры для выполнения сложных как в структурном, так и в вычислительном плане заданий. В частности, обсуждаемое сопряжение обеспечивает доступ к развитой системе вычислительных ресурсов широкому кругу пользователей в научно-исследовательских, правительственных, промышленных и коммерческих организациях для решения задач, требующих массовой обработки большого количества данных, в фундаментальных и прикладных научных исследованиях по медицине, биологии, метеорологии, государственному и муниципальному управлению, а также в исследованиях в области высоких технологий.

Внешними ресурсами могут являться как отдельные вычислительные кластеры или суперкомпьютеры, так и системы распределенных вычислений грид-уровня. Важно только наличие в системе менеджера заданий, обеспечивающего их запуск и мониторинг выполнения. Предлагаемая технология пригодна как для подключения к грид-системе отдельных кластеров или суперкомпьютерных ресурсов, так и для подключения других грид-систем (образования *федераций* грид-систем).

Основное внимание при разработке технологии уделялось следующим требованиям:

- 1) Минимальное вмешательство в работу внешних ресурсов и отказ от установки на них какого-либо ПО, действующего в «фоновом режиме», то есть запускаемого периодически без запроса из задания пользователя или администратора ресурса;
- 2) Максимальное использование возможностей ресурс-менеджера внешнего ресурса с обязательным запуском вычислительных заданий только через него;
- 3) Распределение использования внешних ресурсов между заданиями локальных пользователей и заданиями ГРИД с осуществлением контроля за распределением ресурсов средствами системного администратора внешнего ресурса.

4) Максимальный доступ пользователя системы ГРИД к информации ресурс-менеджера внешнего ресурса в рамках прав доступа, устанавливаемых системным администратором внешнего ресурса.

Суть предлагаемой технологии состоит в использовании для запуска заданий на внешние ресурсы специальных «транзитных» узлов ГРИД-системы. Эти узлы представляют собой специализированные вычислительные ресурсы ГРИД-системы (компьютерные элементы и/или рабочие ноды) на которых устанавливается «транзитный сервис» - ПО, обеспечивающее передачу заданий менеджеру внешнего ресурса, а также мониторинг заданий и доступ к информации о конфигурации ресурса. По-существу «транзитный сервис» является оберткой над клиентской частью ресурс-менеджера, которая устанавливается и специальным образом конфигурируется на транзитном узле.

Для запуска задания или шага задания на внешнем ресурсе через транзитный узел, формируется специальное одношаговое «транзитное» задание ГРИД-системы, направляемое на транзитный узел. «Транзитное» задание содержит обращение к «транзитному сервису» (то есть «транзитный сервис» описывается как исполнимая команда задания), а в качестве аргумента передается описание задачи, выполняемой на внешнем ресурсе, в форме, которая соответствует требованиям менеджера этого ресурса. Например, если в качестве менеджера ресурса используется PBS, то задание описывается в виде PBS-скрипта [1], если внешний ресурс является внешней ГРИД-системой, то описание делается на языке описания заданий этой системы и т.п. Необходимые файлы пользователя (Input Sandbox) передаются на транзитный узел средствами ГРИД-системы, а затем ресурс-менеджеру внешнего ресурса средствами транзитного сервиса.

Базовый уровень технологии не предполагает унификации языка описания заданий для различных типов менеджеров ресурсов, оставляя составление заданий для нужных менеджеров ресурсов на усмотрение пользователя. Это не мешает в дальнейшем при необходимости дополнять ПО транзитного узла средствами автоматической конвертации заданий для конкретного менеджера ресурса на подходящих универсальных языках, удобных для определенных групп пользователей. Применение «не универсального» базового уровня ПО вызвано ориентацией технологии на большую вероятность использования в сложных задачах вычислительных ресурсов уникальной конфигурации, использование возможностей

которой с трудом поддается описанию на универсальных языках. С другой стороны, если некоторая унификация языка описания заданий уместна, ничто не мешает разместить соответствующее ПО на транзитных узлах.

Настоящая работа посвящена развитию разработанной ранее технологии запуска заданий системы glite [2,3] на ресурсах СКИФ-ГРИД [4,5] и представляет собой описание реализации описанной выше архитектуры для интеграции ресурсов ГридННС в систему СКИФ-ГРИД, использующую ПО UNICORE-6 [6]. В качестве менеджера заданий используется система управления заданиями ГридННС (PilotCLI) [7].

Система PilotCLI позволяет управлять многошаговыми заданиями, отдельные шаги которых, называемые задачами, могут выполняться на различных локальных ресурсах (кластерах, суперкомпьютерах и т.п.), входящих в состав ГридННС. Запуск отдельных задач реализуется только как запуск одношагового задания. Существенной особенностью системы является отсутствие поддержки работы с локальными файлами пользователя. Это связано с необходимостью обеспечить возможность отключения пользователя от ГридННС сразу после запуска задания и обработка задания в автономном режиме. Поэтому предполагается, что все необходимые файлы пользователь предварительно размещает на сертифицированных для работы с ГридННС GridFtp-серверах. Соответственно, ППО транзитного узла («транзитный сервис») должно поддерживать взаимодействие с таким сервером. Следует отметить, что ПО для доступа к таким серверам целесообразно размещать и на клиентской машине пользователя, введя его в состав пользовательского интерфейса. Также в состав пользовательского интерфейса необходимо ввести ПО для получения проху-сертификатов, необходимых для авторизации в ГридННС.

Апробацию разрабатываемой технологии предполагается проводить на специальном макете с использованием предварительной версии системы ГридННС, находящейся в настоящее время в стадии разработки.

#### ***Описание макета.***

Макет состоит из двух машин. На первой машине, представляющей модель пользовательского интерфейса пользователя UNICORE, развернута клиентская часть

системы UNICORE-6. В настоящей версии используется командный интерфейс (UCC), но может использоваться и любой другой пользовательский интерфейс UNICORE-6. Кроме того там развернуто ПО, для получения проху-сертификатов РДИГ-СА, пригодных для работы с ГридННС.

На второй машине, являющейся моделью собственно «транзитного узла» развернуты следующие системы:

- 1) Сервер UNICORE-6;
- 2) Командный интерфейс (Pilot-CLI) для запуска заданий системы ГридННС;
- 3) Сервер GridFtp;
- 4) Система авторизации ГридННС, обеспечивающая возможность взаимодействия сервера GridFtp с ГридННС;
- 5) Специализированное ППО для реализации транзитного запуска заданий в ГридННС и контроля за их прохождением.

Развертывание систем, упомянутых в пунктах 4) и 5), связано с особенностями технологии запуска заданий в настоящей версии системы ГридННС, в которой не предусмотрено работы с локальными файлами пользователя. Возможно только использование файлов, доступных GridFtp серверам, сертифицированным для работы с ГридННС.

### ***Структура ППО для транзитного запуска заданий.***

ППО транзитного сервиса состоит из следующего набора скриптов, которые могут использоваться в заданиях пользователя UNICORE (название скрипта при этом будет передаваться с помощью параметра EXECUTABLE). При этом выполнение соответствующих команд предполагает наличие на транзитном узле проху-сертификата пользователя.

Следует подчеркнуть, что работа разрабатываемого ППО не зависит от UNICORE и оно может использоваться для запуска и контроля заданий в ГридННС из любой ГРИД-системы, обеспечивающей доставку на транзитную узел локальных файлов пользователя, содержащих описания задания ГридННС и проху-сертификата.

ППО транзитного узла содержит команды запуска и контроля над задачами, поскольку предполагается, что ресурсы ГридННС могут использоваться пользователем UNICORE (СКИФ-ГРИДа) как для выполнения заданий, так и для выполнения отдельных задач.

***Команды, обеспечивающие контроль за выполнением задания:***

nns-run	запуск задания в гридННС;
nns-task-run	запуск задачи в гридННС;
get-env	окружение в котором выполняется шаг задания UNICORE;
nns-get-status	текущий статус задания;
nns-get-task-status	текущий статус задачи, запущенного из шага задания ГридННС;
nns-get-jobinfo	информация о задании ГридННС;
nns-get-taskinfo	информация о задаче (шаге задания) ГридННС.

***Команды для получения информации о ресурсах ГридННС,  
доступных с помощью данного транзитного узла:***

nns-info	информация о ресурсе, входящем в ГридННС;
nns-resource-list	список ресурсов ГридННС, доступных через данный транзитный узел.

***Команды, обеспечивающие доступ к GridFtp серверу:***

ftp-save	сохранить локальный файл (или набор файлов) на GridFtp сервер транзитного узла;
ftp-load	загрузить локальный файл (файлы) с GridFtp сервера транзитного узла.

Следует отметить, что в отличие от обычных команд для работы с GridFtp, эти утилиты знают «адрес» соответствующего сервера.

### *Заключение.*

В настоящее время описываемая система находится в стадии разработки. Соответственно, набор команд, входящих в «транзитный сервис», будет дополняться, а функциональность уже существующих команд расширяться. Кроме того, чтобы упростить описание заданий UNICORE, осуществляющих обращение к транзитным сервисам, вероятно будет разработано специализированное ПО, располагаемое на пользовательском интерфейсе пользователя UNICORE.

Работа выполнена в соответствии с договором 2009-СГ-06 при финансовой поддержке Федерального агентства по науке и инновациям (г/к 01.647.11.2004 и 02.740.11.0388).

### ЛИТЕРАТУРА

1. <http://www.clusterresources.com/pages/products/torque-resource-manager.php>
2. <http://glite.web.cern.ch/glite/>
3. <https://edms.cern.ch/file/722398//gLite-3-UserGuide.pdf>
4. V. Edneral, V. Plyin, A. Kryukov, G. Shpiz, L. Shamardin. *Implementation of job submission interface from EGEE/WLCG GRID infrastructure to SKIF series supercomputers.*  
Proceedings of Conference GRID 2008 (Dubna, June 30 - July 4, 2008), pp. 179 – 181.
5. Бережной А.В., Еднерал В.Ф., Ильин В.А., Крюков А.П., Шпиз Г.Б., Шамардин Л.В.,  
*Интерфейс между EGEE/LCG гридом и суперкомпьютерами серии СКИФ.*  
Препринт НИИЯФ МГУ № 2008-8/844 (2008), 6 стр.
6. <http://www.unicore.eu>
7. <https://ngrid.ru>



**Александр Викторович Бережной, Виктор Федорович Еднерал,  
Вячеслав Анатольевич Ильин, Александр Павлович Крюков,  
Григорий Борисович Шпиз**

**ТЕХНОЛОГИЯ ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ГРИДННС  
ПОЛЬЗОВАТЕЛЯМИ СИСТЕМЫ СКИФ-ГРИД**

Препринт НИИЯФ МГУ № 2009-10/854

Работа поступила в ОНТИ 24 ноября 2009