

**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, Л.В. Шамардин

**ИНТЕРФЕЙС МЕЖДУ EGEE/LCG ГРИДОМ И
СУПЕРКОМПЬЮТЕРАМИ СЕРИИ СКИФ**

Препринт НИИЯФ МГУ № 2008-8/844

Москва, 2008 г.

УДК

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

e-mail: aber@trtk.ru

**ИНТЕРФЕЙС МЕЖДУ EGEE/LCG ГРИДОМ
И СУПЕРКОМПЬЮТЕРАМИ СЕРИИ СКИФ**

Препринт НИИЯФ МГУ № 2008-8/844

Аннотация

Разработан и протестирован интерфейс между EGEE/LCG ГРИДОМ и СКИФ-кластером. Настоящий интерфейс позволяет использовать ресурсы суперкомпьютеров СКИФ в рамках стандартной системы запуска, мониторингования и получения результатов ГРИД-заданий (gLite).

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

**INTERFACE BETWEEN EGEE/LCG GRID AND
SUPERCOMPUTERS OF THE SKIF TYPE**

Preprint MSU SINP N 200-8/844

Abstract

The interface between EGEE/LCG GRID and the SKIF cluster has been developed and tested. This interface allows to use the resources of SKIF supercomputers within the standard system for GRID jobs submitting, monitoring and getting GRID jobs results (gLite).

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

© НИИЯФ МГУ, 2008, <http://www.sinp.msu.ru>

Сопряжение ГРИД-инфраструктуры РДИГ/EGEE/LCG со СКИФ-ресурсами является важным элементом программно-технического комплекса, предназначенного для доступа через ГРИД-инфраструктуру РДИГ к СКИФ-ресурсам в научно-исследовательских, правительственных, промышленных и коммерческих организациях для решения задач, требующих массовой обработки большого количества данных в фундаментальных и прикладных научных исследованиях, медицине и биологии, метеорологии, в государственном и муниципальном управлении, а также на предприятиях сектора высоких технологий.

Основное внимание при решении поставленной задачи уделялось требованию минимального вмешательства в работу самих кластеров СКИФ и их программное обеспечение, для чего была выбрана схема, в которой взаимодействие компонентов ГРИДа с кластером СКИФ происходит с помощью специального шлюза, построенного на базе рабочего узла (WN) ППО gLite [1]. При этом на вычислительных машинах кластера устанавливается клиентская часть системы управления заданиями. В нашем конкретном случае использовалась система Torque [2].

Предложенное решение обеспечивает не только возможность запуска заданий из EGEE/WLCG, но также прозрачность информационных систем СКИФ кластера для пользователей и ресурс-брокеров ГРИД. Для запуска через ГРИД задание описывается в обычном JDL файле [3] и запускается с машины со стандартного пользовательского интерфейса.

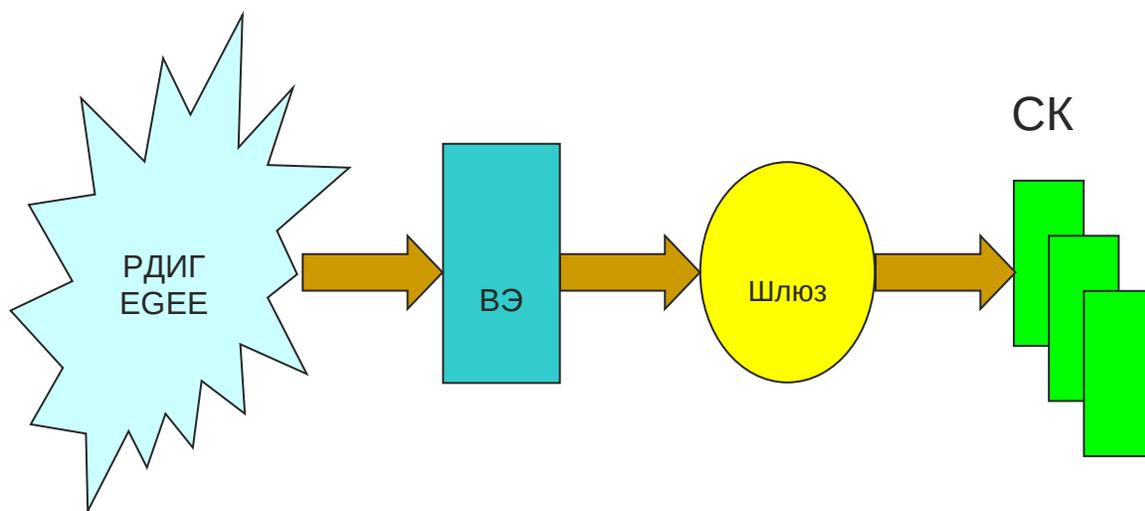


Рис. 1. Упрощенная схема узла СКИФ-ГРИДа: ВЭ – вычислительный элемент (CE gLite), Шлюз на основе WN gLite, осуществляющий передачу РДИГ-заданий в СК СКИФ.

Таким образом, разработанная нами схема (см. также [4]) позволяет «развязать» ППО РДИГа и суперкомпьютера СКИФ (см. рис. 1).

Созданный на основе этой схемы макет (testbed) узла СКИФ-ГРИДа состоит из следующих основных компонент:

- ❖ Вычислительного элемента (CE gLite);
- ❖ Шлюза (WN gLite) между компонентами EGEE и моделью суперкомпьютера СКИФ;
- ❖ макета суперкомпьютера СКИФ на основе 16 – ядерного MPI-кластера в составе:
 - 4-х узлов по 2 двухъядерному CPU в каждом (процессоры Xeon 5100 с рабочей частотой 3 GHz);
 - 1 гигабайта RAM на каждом узле;
 - сетевого соединения – Infiniband.

При этом на СКИФ-ГРИД сайте установлено ППО:

- ❖ на Вычислительном элементе и Шлюзе:
 - ОС Scientific Linux 4;
 - ППО gLite 3 для CE и WN. В том числе PBS сервер и менеджер очередей (Maui) на CE и Torque (or PBS) клиент на WN;
 - разработанный набор bash скриптов для передачи заданий на СК СКИФ;
- ❖ на макете суперкомпьютера СКИФ:
 - Torque (or PBS) клиент;
 - MPICH (или MPICH-2) библиотека, версия под Infiniband.

Для успешной установки ППО на узел СКИФ ГРИДа необходимо обеспечить следующие условия:

- ❖ беспарольное ssh, rsh взаимодействие между узлом и СКИФ-кластером для его локальных пользователей (например для VO dteam это dteam001, dteam002, ...). В частности, все эти пользователи должны быть известны на вычислительных элементах СКИФ-кластера;
- ❖ файловая система должна быть смонтирована таким образом, чтобы директория “/home”

на СКИФ-кластере и на других компьютерах узла была общей;

- ❖ возможность запуска заданий через PBS-сервер на вычислительном элементе (ВЭ) на машины СКИФ-кластера, где установлены Torque (or PBS)-клиенты;
- ❖ необходимо обеспечить **невозможность** прямого запуска EGEE/РДИГ задач на машинах СКИФ-кластера. Они должны стартовать только на Шлюзе.

В настоящий момент нами разработан и протестирован СКИФ-ГРИД узел, способный работать реальной в EGEE/РДИГ среде и позволяющий использовать стандартную систему запуска, мониторингования и получения результатов ГРИД-заданий (gLite). Разработанное ППО снабжено подробной документацией, описывающей как процесс установки, так и процесс эксплуатации СКИФ-ГРИД узла.

Работа выполнена по контракту Союзного государства России и Беларуси № СГ-2/07 и поддержана грантом РФФИ № 08-07-00294.

ЛИТЕРАТУРА

1. <http://glite.web.cern.ch/glite/>
2. <http://www.clusterresources.com/pages/products/torque-resource-manager.php>
3. <https://edms.cern.ch/file/722398/gLite-3-UserGuide.pdf>
4. V. Edneral, V. Ilyin, 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).

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

**ИНТЕРФЕЙС МЕЖДУ EGEE/LCG ГРИДОМ И
СУПЕРКОМПЬЮТЕРАМИ СЕРИИ СКИФ**

Препринт НИИЯФ МГУ № 2008-8/844

Работа поступила в ОНТИ *19 ноября 2008*