

Организация взаимодействия с различными системами управления кластерами в рамках программного комплекса для квантово-механических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов

В.Д. Кустикова, А.В. Линева, И.Н. Лозгачев, А.В. Сенин, А.В. Сысоев

Нижегородский госуниверситет им. Н.И. Лобачевского

Введение

В 2008-2009 гг. Нижегородский государственный университет им. Н.И. Лобачевского (ННГУ) участвует в реализации проекта по разработке высокопроизводительного программного комплекса для квантово-механических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов (головным исполнителем является Санкт-Петербургский государственный университет информационных технологий, механики и оптики).

Цель проекта состоит в создании программного комплекса прикладных алгоритмов, программ и математических моделей для квантово-механических расчетов: от геометрических характеристик сложной атомно-молекулярной структуры до макропараметров системы. В данной работе рассматривается вопрос сопряжения программного компонента управления платформами исполнения с системами управления кластером.

Взаимодействие с различными системами управления кластерами

Основная задача программного компонента управления платформами исполнения состоит в предоставлении программного интерфейса управления кластерами, независимого от типа используемых систем управления. Непосредственное взаимодействие является задачей программного модуля «Трансивер», который должен обеспечить сопряжение со следующими системами управления: Torque [1], Microsoft Windows HPC Server 2008 [2], «Метакластер» [3] (разработка ННГУ).

Трансивер обеспечивает выполнение следующих операций:

- предоставление корректной информации о составе (имена узлов кластера) и аппаратных и программных характеристиках (размер жесткого диска, количество ядер, частота процессора, список установленного программного обеспечения и т.д.) кластера;
- предоставление корректной информации о текущей загрузке кластера (текущая загрузка процессора каждого узла, загрузка сетевого интерфейса, идентификаторы задач, выполняющихся на каждом узле, и т.п.);
- запуск и остановка заданий;
- резервирование ресурсов для запуска задачи и отмена резервирования.

Механизмы взаимодействия

Разработчики систем управления кластерами предоставляют различные наборы средств взаимодействия, посредством которых можно реализовать перечисленную выше функциональность. Трансивер использует различные возможности интерфейсов для сопряжения с различными системами управления:

- трансивер Torque использует удаленный вызов команд интерфейса системы управления средствами удаленного терминала (ssh, [5]);
- трансивер Windows HPC Server 2008 использует программный интерфейс .NET API, предоставляемый компанией Microsoft, и командный интерфейс Windows PowerShell;
- трансивер Метакластера использует интерфейс .NET Remoting, позволяющий получить удаленный доступ к системе управления кластером «Метакластер».

Принцип работы трансивера

Трансивер представляет собой Windows service или Linux daemon и является сервером распределенного приложения, реализованным с применением технологии .NET Remoting. Основная задача трансивера – обслуживание внешних запросов:

- запрос статических и динамических характеристик кластера – обрабатывается посредством обращения к системе управления кластером, либо к системе мониторинга Ganglia в зависимости от возможностей системы управления (рис. 1);
- запросы на запуск и остановку задачи, резервирование всех свободных ресурсов для запуска задачи – обрабатываются посредством обращения к системе управления кластером.

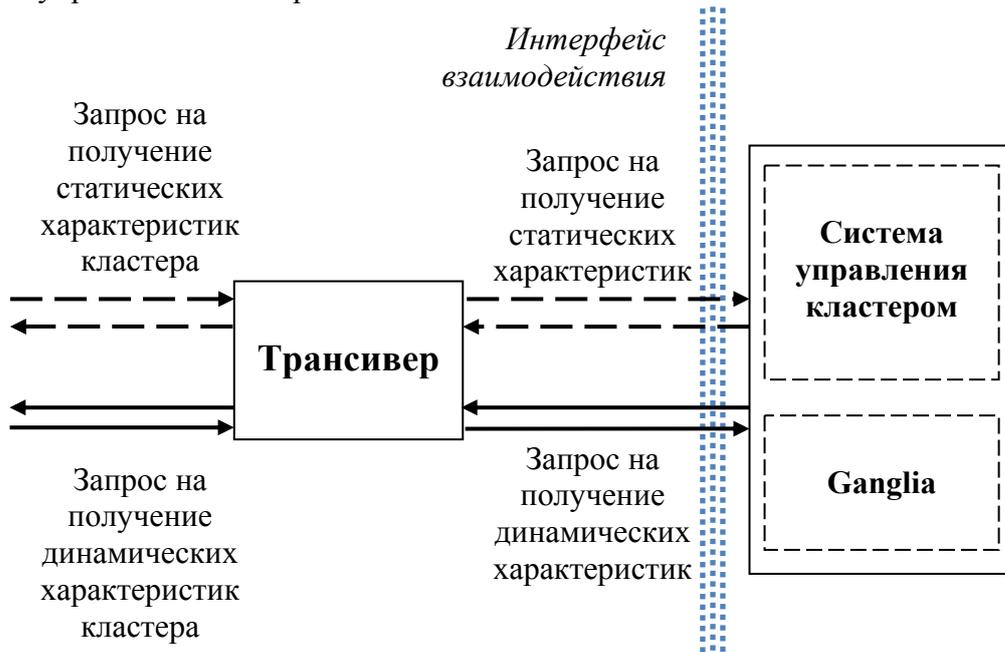


Рисунок 1. Получение статических и динамических характеристик кластера.

Рассмотрим подробнее обслуживание трансивером важнейшей операций: выполнение и обработка завершения выполнения задачи (см. рис.2). Типовое обслуживание данных операций реализуется посредством выполнения следующей последовательности действий:

- прием запроса на выполнение задачи, содержащий ее описание (рис. 2, п.1);
- копирование входных данных задачи с FTP-сервера на локальный файловый сервер кластера (рис. 2, п.2);
- формирование запроса на исполнение задачи к системе управления кластером в соответствии с выбранным интерфейсом взаимодействия (рис. 2, п.3);
- прием обратного вызова, который может выполнять система управления кластером после завершения задачи (рис. 2, п.4); если система управления не поддерживает обратный вызов – периодический опрос состояния выполняющихся задач;
- копирование выходных данных задачи с локального файлового сервера кластера на FTP-сервер (рис. 2, п.5);
- выполнение обратного вызова Интегратора управления – модуля, приславшего запрос на исполнение задачи (рис. 2, п.6).

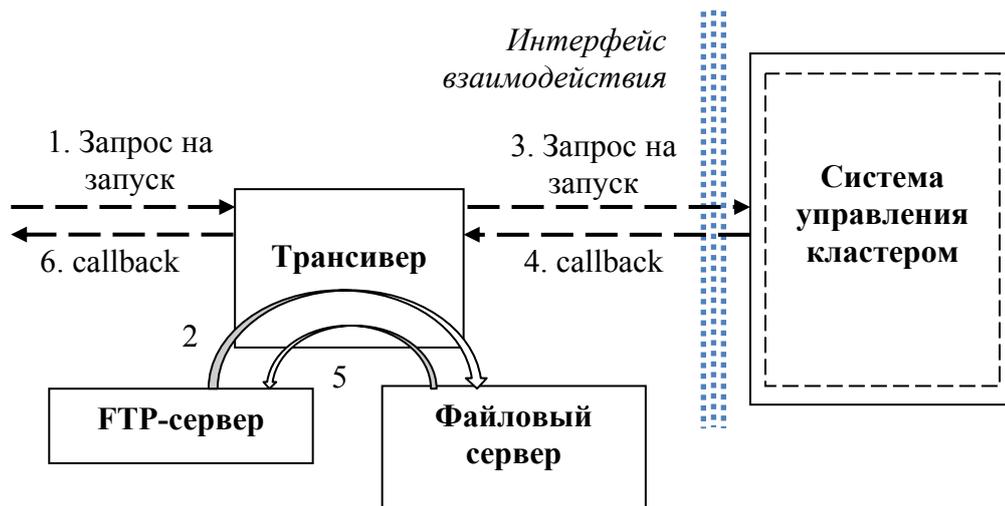


Рисунок 2. Схема выполнения запроса на выполнение задачи.

Для обеспечения параллельного обслуживания запросов используется многопоточная реализация трансивера. Дополнительные потоки выполняют опрос текущего состояния задач, контроля состояния задачи средствами, независимыми от системы управления кластером, копирование выходных данных задач и удаление завершившихся задач.

Заключение

Архитектура программного компонента управления платформами исполнения позволяет выполнять сопряжение с различными системами управления кластером посредством использования соответствующих трансиверов. К настоящему моменту разработаны трансиверы для систем управления Torque, Microsoft Windows HPC Server 2008, «Метакластер». Текущие версии трансиверов развернуты на территориально распределенной тестовой базе, включающей 5 кластеров с общим количеством узлов 96.

Исследования выполнены в рамках ОКР по теме «Разработка высокопроизводительного программного комплекса для квантово-механических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов».

Литература

1. Официальная страница системы управления Torque – [<http://www.clusterresources.com/products/torque-resource-manager.php>].
2. Официальная страница Microsoft HPC Server 2008 – [<http://www.microsoft.com/hpc/en/us/>].
3. Система управления «Метакластер» – [www.cluster.software.unn.ru].
4. В.Н. Васильев, А.В. Бухановский, С.А. Козлов, В.Г. Маслов, Н.Н. Розанов. Высокопроизводительный программный комплекс моделирования наноразмерных атомно-молекулярных систем // Научно-технический вестник. СПбГУ ИТМО «Технологии высокопроизводительных вычислений и компьютерного моделирования», № 54. СПб: Университетские телекоммуникации, 2008. С.3-13.
5. Библиотека SecureShell для .NET – [<http://www.tamirgal.com/blog/page/SharpSSH.aspx>].