

УДК 681.324

М.А. Волк

Харьковский национальный университет радиоэлектроники, Харьков

СТРУКТУРА ПРОГРАММНОГО КОМПЛЕКСА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ЭЛЕМЕНТОВ GRID-СИСТЕМ ДЛЯ НАУЧНЫХ ИССЛЕДОВАНИЙ

В работе приводится структура разработанной программной имитационной модели GRID - системы, рассматриваются ее элементы, анализируются пути использования модели. Модульный характер модели дает возможность ее использования в научных исследованиях, связанных с разработкой новых методов поведения отдельных элементов GRID - систем. Модель может быть также полезна специалистам в области распределенного имитационного моделирования и построения параллельных, многопоточных и распределенных программных систем.

Ключевые слова: GRID-система, имитационная модель, программный комплекс.

Введение

В настоящее время в Украине ведется формирование Национальной GRID-инфраструктуры, которая должна стать частью межнациональной системы, открывающей пользователю доступ к разнообразным ресурсам [1]. Среди этих ресурсов – огромные базы данных, локальные и глобальные сети, кластера. Все перечисленные ресурсы для организации логической связи между собой используют промежуточное программное обеспечение (middleware), которое само по себе представляет сложную программную систему, состоящую из десятков программ, реализующих различные сервисы.

Параллельно с решением технических проблем по созданию и отладке такого программного обеспечения постоянно ведутся научно-исследовательские работы по оптимизации его функционирования [2]. Действительно, во владении GRIDa оказались большие и дорогостоящие ресурсы научных учреждений и других организаций. Если не уделять внимание вопросам эффективности использования таких ресурсов, могут возникнуть проблемы, связанные со значительными потерями в стоимости решения задач, надежности, безопасности, пропускной способности и т.д.

К наиболее часто поднимаемым вопросам, связанным с исследованиями в построении GRID систем, относятся:

- распределение ресурсов;
- эффективное использование трафика;
- управление очередью заданий;
- информационная безопасность;
- надежность и т.п.

Каждый из этих вопросов требует проведения отдельных научных и практических исследований. Однако последние, особенно в части проведения экспериментальной работы, испытывают значительные трудности как раз в связи с самой природой

GRID как распределенного программного обеспечения [3].

В Харьковском национальном университете радиоэлектроники ведутся работы по созданию распределенной имитационной модели GRID, которая позволит реализовать и исследовать различные методы и алгоритмы определенных элементов GRID. В данной статье приводится структурная схема разрабатываемого программного обеспечения и анализируются пути его использования.

Структура программного комплекса

Программное обеспечение имеет модульный характер, изначально ориентированный на распределение между вычислительными ресурсами. Так как для обмена данными между модулями используются сетевые подходы, оно может функционировать как на одном компьютере (для небольших вычислительных задач), так и в локальной, глобальной сети ЭВМ или кластере. Таким образом, масштаб задач, которые можно исследовать при помощи предлагаемой имитационной модели, может быть достаточно большим.

Остановимся на описании модулей, входящих в состав разрабатываемого программного обеспечения (рис. 1).

1. Интерфейс пользователя. Представляет собой набор программ визуального интерфейса, позволяющий поставить эксперимент. Пользователь может задать основные количественные и качественные параметры задания, потоков заданий, конфигурацию GRID-системы и оценить результаты эксперимента.

2. Постановщик задания. Принимает реальное или виртуальное задание для исполнения в GRID. Осуществляет связь с middleware GRID для передачи задания в виртуальную организацию [4].

3. Блок имитации потока заявок. Программа, размещаемая на удаленном компьютере, моделирую-

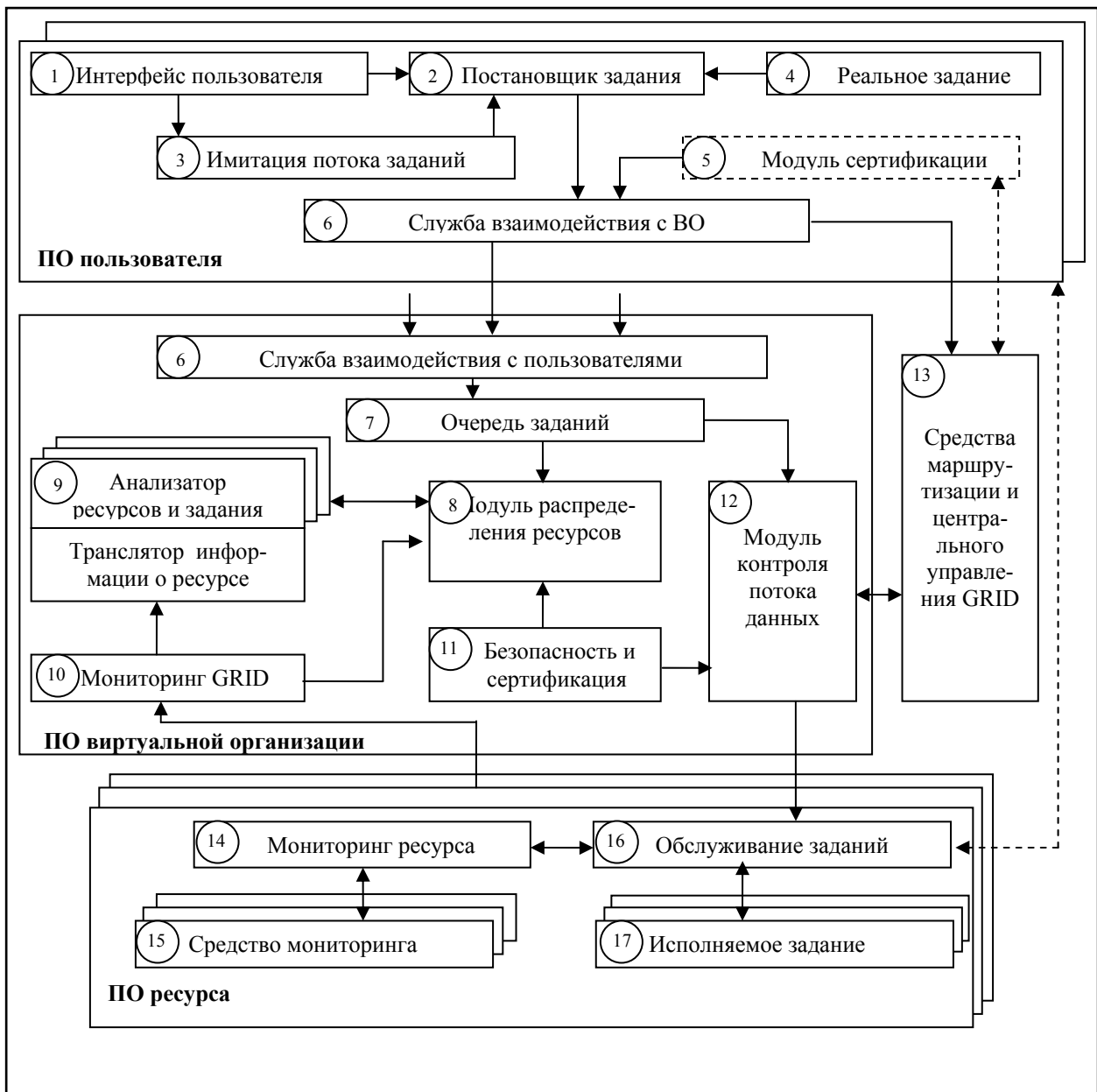


Рис. 1. Структура программного комплекса моделирования GRID

щая потоки заявок в GRID. Пользователь имеет возможность задать параметры задания:

- количество программ в задании;
- размер программы (программ);
- размер передаваемых исходных данных (потока данных);
- размер передаваемых данных результата (потока данных);
- объемы трафика во время параллельного исполнения программ.

Кроме этого, существует возможность управлять потоком заданий, изменяя интервалы времени поступления заданий, их количество и другие характеристики.

Для реализации этих возможностей разработано следующее:

а) формат представления заданий (основанный на языках представления заданий в GRID) [5];

б) модуль, реализующий законы распределения (равномерный, экспоненциальный, нормальный и др.) на основе Boost.Random [7].

4. Система может принять в качестве заявки реальное задание. В этом случае из заявки выделяются три части: исполняемые файлы, файлы исходных данных и файл-описание задания на языке jdl (или ему подобном) [5]. Файл-описание может использоваться отдельно для исследований алгоритмов и стратегий планирования распределения ресурсов в GRID [6]. Целью эксперимента также может быть прогнозирование поведения различных конфигураций GRID при постановке данного задания.

5. Модуль сертификации определяет права конкретного пользователя на множество доступных ресурсов.

6. Служба взаимодействия с ВО. Программное обеспечение, поддерживающее сетевой обмен данных пользователя с виртуальной организацией.

7. Блок очереди заданий. Обеспечивает получение задания (или информации о нем) и размещение его в очереди к блоку распределения заявок.

8. Модуль распределения ресурсов. Совместно с программным обеспечением, приведенным в п.9, обеспечивает распределение заявок среди доступных ресурсов на основе данных о задаче и ресурсах. Должен поддерживать подключение разных алгоритмов распределения. В дальнейшем – поддерживать выбор нужного алгоритма и динамическую смену алгоритма при изменении параметров задачи или ресурсов.

9. Анализатор ресурсов и задания. Содержит алгоритм распределения ресурсов на основе информации о заданиях и доступных ресурсах. Так как алгоритмы распределения используют разные входные данные, для обеспечения их работы требуется наличие транслятора информации о ресурсе.

10. Блок мониторинга GRID. Сканирует доступные ресурсы, создает динамические таблицы характеристик доступных ресурсов и каналов данных:

- быстродействие процессора;
- количество ядер;
- средняя загрузка процессора;
- объемы кеш, оперативной, виртуальной, внешней памяти (максимальный, минимальный, средний);
- характеристики сети (скорость обмена данными, загрузка канала – максимальная, средняя, минимальная) и т.п.

11. Модуль безопасности и сертификации. Определяет ресурсы, доступные конкретному пользователю на основе выданных им сертификатов.

12. Модуль контроля потока данных. Управляет процессом размещения задач на ресурсах, передачей данных заданий, синхронизацией задач во время выполнения, устранением ошибок, связанных с изменением конфигурации и состава ресурсов.

13. Средства маршрутизации и центрального управления GRID. Имитируют поведение транспортных средств GRID по передачи заданий и данных между пользователем и ресурсами.

14. Мониторинг ресурса. Программа, обеспечивающая передачу данных о ресурсе, получаемых от программ мониторинга (п.15) в виртуальную организацию (п.10).

15. Средство мониторинга – представляет собой программу (программы), собирающую разнородную информацию о ресурсе.

16. Исполнитель подзадачи. Размещается на конкретном ресурсе и обеспечивает принятие на

исполнение задания и данных, а также запуск задания, поддержку его выполнения, передачу результатов исполнения.

17. Исполняемое задание. Каждый модуль выполнен в виде отдельной программы. Протоколы обмена данными между модулями основаны на открытом стандарте *.xml, что позволяет просматривать потоки данных между любыми модулями. Кроме того, данная особенность дает возможность использования модулей выборочно согласно задачам исследования.

Возможные варианты исследований

В предыдущем разделе показано, что имитационную модель GRID системы можно использовать как целиком, для анализа работы системы в целом, так и "помодульно", для постановки экспериментов с отдельными подсистемами. В данном разделе приведем некоторые примеры, которые дают представление о возможностях исследований для научных сотрудников, работающих в области построения сложных вычислительных систем.

Для исследователей в области систем массового обслуживания возможно проведение экспериментов по анализу математических моделей потоков заданий, управлению очередями, определению нагрузочных характеристик отдельных модулей, моделей предоставления трафика и т.п.

Особый класс подзадач, активно исследуемый в настоящее время – методы распределения ресурсов в GRID-системах. Это важный вопрос как для владельца сегмента GRID, предоставляющего свои ресурсы другим пользователям, так и для пользователя, который эффективность решения задачи рассматривает в терминах время-стоимость. Для апробации алгоритмов распределения заявок создан специальный открытый модуль, который в режиме плагинов (plug-in) дает возможность подключения к системе своих оригинальных подпрограмм. Таким образом, исследователь может оценить эффективность предложенных алгоритмов на различных классах задач, сравнить их с существующими методами, а в дальнейшем встроить разработанный программный модель в реальную GRID систему.

Подобным же образом разработана подсистема мониторинга ресурсов. Конкретный пользователь может воспользоваться уже разработанным модулем мониторинга. Если необходимо использовать дополнительные характеристики ресурса, не заложенные в базовый вариант программного обеспечения, можно дополнить систему мониторинга своей подпрограммой, которая реализует сбор и интерпретацию нового параметра ресурса. Существует возможность добавить подобную функцию и на стороне подсистемы распределения ресурсов.

GRID-система разрабатывается с учетом боль-

ших объемов данных, которые должны передаваться между ее сегментами и пользователями. Поэтому важным направлением исследований является изучение вопросов оптимизации и обеспечения целостности передачи данных в сетях. И здесь имитационное моделирование дает возможность проведения множества экспериментов, которые либо невозможны, либо нежелательны в реальных системах. В качестве отдельного класса задач, решаемых с использованием создаваемой модели, можно выделить создание методов и средств аутентификации, сертификации и управления политикой безопасности программного обеспечения и данных в больших информационных системах, одной из которых и является GRID.

Одним из целевых направлений использования GRID-систем является проведение с их помощью имитационного моделирования больших систем. Разработанное программное обеспечение представляет интерес также с точки зрения построения систем распределенного имитационного моделирования [3].

Выше приведено только несколько вариантов использования разрабатываемой модели. Этот список может быть расширен за счет задач обеспечения безопасности информации, сертификации, взаимодействия нескольких виртуальных организаций и т.п. Важным является сама возможность испытания разработанного модуля в реальных условиях на большом количестве тестовых задач.

Выводы

Анализируя научные работы, выполняемые в области разработки GRID-систем, можно сделать вывод о том, что в большинстве из них практическим результатом является имитационное моделирование отдельно взятого элемента системы. Программные средства, описанные в данной работе, позволяют широкому кругу исследователей апробировать свои предложения на имитационной модели, приближенной к реальной системе, исследовать их в различных режимах, например, близких к критическим. После отладки и модификации, разработанный

модуль может быть легко перемещен в реальную GRID систему. Разработанная модель может быть также полезна специалистам в области распределенного имитационного моделирования и построения распределенных программных систем.

В дальнейшем планируется интегрировать элементы разработанной программной имитационной модели в middleware GRID для проведения более масштабных экспериментов.

Список литературы

1. Петренко А.И. GRID технології в науці і освіті / А.И. Петренко // Матеріали 9-ої Міжнародної конференції «Системний аналіз та інформаційні технології», Київ, 2007. – С.138-140.
2. Aida K. Performance Evaluation Model for Scheduling in a Global Computing System / K. Aida, A. Takefusa, H. Nakada, S. Matsuoka, S. Sekiguchi, U. Nagashima // Int. J. of High Performance Computing Applications, 14(3), 2000. – P. 268-279.
3. Волк М.А. Структурная организация поведенческого имитационного моделирования в GRID / М.А. Волк // Системи обробки інформації. – X.: XV ПС, 2007. – Вып. 9(67). – С. 41-45.
4. Foster I. The Anatomy of the Grid: Enabling Scalable Virtual Organizations / I. Foster, C. Kesselman, S. Tuecke // International Journal of High Performance Computing Applications. – 2001. – 15 (3). – P. 200-222.
5. WPI - WMS Software Administrator and User Guide job description language howto -. DataGrid-01-TEN-0118-0_0-December 17, 2001, Rome. [Электронный ресурс]. – Режим доступа к ресурсу: http://www.infn.it/workload-grid/docs/DataGrid-01-TEN-0118-0_document.pdf.
6. Методы и модели планирования ресурсов в GRID-системах: монография / В.С. Пономаренко, С.В. Лустровой, С.В. Минухин, С.В. Знахур. – X.: ВД „ІНЖЕК”, 2008. – 408 с.
7. Abrahams D. C++ Template Metaprogramming: Concepts, Tools, and Techniques from Boost and Beyond – Addison Wesley Professional / D. Abrahams, A. Gurtovoy. – 2004. – 312 с.

Поступила в редколлегию 24.04.2009

Рецензент: д-р техн. наук, проф. О.Г. Руденко, Харьковский национальный университет радиоэлектроники, Харьков.

СТРУКТУРА ПРОГРАМНОГО КОМПЛЕКСУ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ ЕЛЕМЕНТІВ GRID-СИСТЕМ ДЛЯ НАУКОВИХ ДОСЛІДЖЕНЬ

М.О. Волк

В роботі наводиться структура розробленої програмної імітаційної моделі GRID-системи, розглядаються її елементи, аналізуються шляхи використання моделі. Модульний характер моделі дає можливість її використання в наукових дослідженнях, пов'язаних з розробкою нових методів поведінки окремих елементів GRID-систем. Модель може використовуватись спеціалістами в галузі розподіленого імітаційного моделювання і побудови паралельних, багатопотокових і розподілених програмних систем.

Ключові слова: GRID-система, імітаційна модель, програмний комплекс.

STRUCTURE OF THE SIMULATION PROGRAM COMPLEX OF GRID-SYSTEMS UNITS FOR SCIENTIFIC RESEARCHES

M.A. Volk

In article the structure over of the developed program simulation model of GRID-systems is brought are the systems, its elements are examined, the ways of the use of model are analysed. Module character of model is given by possibility of its use in scientific researches, related to development of new methods of conduct of separate elements of GRID - systems. The model can be useful also to experts in the field of the distributed simulation and construction of the distributed program systems

Keywords: GRID-system, simulation model, programmatic complex.