

Е. Е. Ивашко

Desktop Grid корпоративного уровня¹

АННОТАЦИЯ. В работе рассказывается о системах Desktop Grid корпоративного уровня, их возможностях, отличительных особенностях и приложениях. Даны примеры реализации систем промежуточного программного обеспечения, предназначенных для организации корпоративных Desktop Grid.

Ключевые слова и фразы: Desktop Grid, Enterprise Desktop Grid, распределенные вычисления

Введение

Высокопроизводительные вычисления являются основным инструментом современных материаловедения, химии, разработки лекарственных препаратов, конструирования высокотехнологичных машин и оборудования, проведения фундаментальных и прикладных исследований. Вычислительные эксперименты позволяют снизить количество необходимых для подтверждения свойств объекта натуральных испытаний, а иногда и вовсе отказаться от их проведения.

Основной фокус исследований и разработок в области высокопроизводительных вычислений направлен на повышение производительности расчетов (создание все более производительного аппаратного и прикладного и системного программного обеспечения), разработку новых математических моделей для решения фундаментальных и прикладных задач, а также организацию процесса вычислений (балансировка нагрузки, планирование задач, защита от сбоев и др.).

¹ Работа поддержана грантами РФФИ 12-07-31147 и 13-07-00008

Как правило, для организации высокопроизводительного моделирования используются специализированные вычислители — вычислительные кластеры. Они обеспечивают большую производительность, хорошо подходят для решения широкого класса ресурсоемких задач, однако требуют больших инвестиций в инфраструктуру и оборудование, а также сложны в обслуживании.

Грид кластеров позволяет, с одной стороны, повысить пиковую производительность всей вычислительной системы, а с другой — распределить затраты между несколькими центрами высокопроизводительных вычислений [1]. При этом существенно сужается класс решаемых задач и возникают проблемы синхронизации, взаимодействия и другие, связанные с функционированием распределенной вычислительной сети.

1. Desktop Grid

Desktop Grid — это объединение в качестве единого логического «суперкомпьютера» большого количества неспециализированных вычислителей (офисных рабочих и персональных компьютеров, ноутбуков и даже смартфонов) относительно невысокой производительности. Такой подход еще больше сужает класс доступных для эффективного решения задач. При этом основные преимущества технологии при соответствующей реализации — практически неограниченные масштабируемость и, следовательно, пиковая производительность грид-системы, устойчивость к сбоям и минимальная стоимость создания и сопровождения вычислительной сети.

С ростом скоростей доступа в Интернет и локальных сетей передачи данных, ростом производительности персональных компьютеров, повышается и актуальность реализации вычислительных проектов на базе Desktop Grid. На текущий момент действует 94 крупных проекта на базе Desktop Grid, объединяя более 200 тысяч активных пользователей, предоставляющих ресурсы более 10 миллионов компьютеров суммарной пиковой производительностью бо-

лее 7 Петафлопс (что соответствует 6-7 месту в списке TOP 500 мощнейших суперкомпьютеров мира).

Технология организации распределенных вычислений Desktop Grid имеет целый ряд особенностей по сравнению с кластерами и грид-кластерами, такими, например, как невозможность гарантировать доступность вычислительного узла, большая гетерогенность и разброс в производительности и скорости доступа к узлам, а также многие другие. Такие особенности требуют повышенного внимания к вопросам управления ресурсами, планированию заданий, обработки ошибок, обеспечению безопасности и др.

Одной из наиболее популярных платформ организации распределенных вычислений в Desktop Grid является BOINC (Berkeley Open Infrastructure for Network Computing) [2]. Из 94 действующих проектов добровольных вычислений 77 основаны на BOINC, т.е. данная платформа де-факто является стандартом для добровольных вычислений [3]. Ярким примером BOINC-проекта в сети Интернет, использующим свободные вычислительные ресурсы компьютеров добровольцев, является проект SETI@home, цель которого — анализ космических радиосигналов. Средняя производительность вычислений этого проекта по состоянию на 11 января 2014 года составляет 1.6 петафлопс [4]. Аналогичные показатели пиковой производительности имеют суперкомпьютеры, находящиеся во втором десятке месте в списке TOP-500 (по состоянию на ноябрь 2013 года). При этом в проекте SETI@home участвует более 1.4 млн. пользователей, а виртуальная вычислительная система состоит из 3.5 млн. узлов.

Платформа BOINC является активно развивающимся открытым (Open Source) программным обеспечением и предоставляет богатую функциональность. При этом она отличается простотой в установке, настройке и администрировании, а также обладает хорошими возможностями по масштабируемости, простоте подключения новых вычислительных узлов, использованию дополнительного программного обеспечения, интеграции с другими грид-системами и т. д.

Помимо BOINC существует целый ряд других разработок, основанных на тех же идеях: X-Com [5], Condor [6] и другие.

2. Desktop Grid корпоративного уровня

Одно из перспективных направлений создания Desktop Grid — это реализация технологии распределенных вычислений в рамках локальных коммуникационных сетей предприятий и организаций, т. е. Desktop Grid корпоративного уровня или Enterprise Desktop Grid. Такой подход позволяет частной компании без существенных инвестиций в оборудование и инфраструктуру получить необходимые вычислительные мощности.

Согласно [7], корпоративная Desktop Grid должна удовлетворять следующим требованиям:

- Эффективность — использование максимального объема свободных ресурсов.
- Устойчивость к сбоям — вычисления должны завершаться с предсказуемой производительностью, скрывая сбои в вычислениях.
- Безопасность — защита целостности вычислений (приложения и данных).
- Масштабируемость — работа при изменении (уменьшении/увеличении) доступных ресурсов с предсказуемым эффектом.
- Управляемость — затраты на администрирование не должны расти с той же скоростью, что и размер сети.
- «Прозрачность» — использование ресурсов не должно отражаться на работе корпоративных приложений.
- Открытость/простота интеграции приложений — запуск приложений с минимальными усилиями.

Desktop Grid корпоративного уровня по своим возможностям находится между сервисными Grid и классическими Desktop Grid, расширяя класс доступных приложений (благодаря повышению скорости передачи данных и возможности обмена большими объемами данных), снижая гетерогенность и вероятность отказа вычислительных узлов, а также повышая управляемость самой сети.

Одной из первых реализаций, ориентированных именно на организацию вычислений в рамках локальной сети предприятия, является система Entropia [7], разрабатывавшаяся в начале 2000-х годов. Другими примерами систем организации Enterprise Desktop Grid являются Alchemi [8] и, созданная позже теми же разработчиками с учетом полученного ранее опыта, Aneka [9]. В работе [10] приведен ряд особенностей и классификация реализаций промежуточного программного обеспечения Desktop Grid, в том числе ориентированных на работу в рамках локальной сети организации.

Примерами перспективных для Enterprise Desktop Grid приложений являются, например, следующие [11]–[14]:

- аналитическая обработка больших наборов данных;
- распределенный рендеринг для студий анимации;
- моделирование промышленных продуктов и имитационное моделирование;
- финансово-экономическое моделирование;
- и другие.

В рамках научной организации Desktop Grid может служить дополнением или — в некоторых случаях — заменой вычислительного кластера [14].

Заключение

Программные системы по организации высокопроизводительных вычислений на базе Desktop Grid вызывают все больший интерес благодаря развитию вычислительной техники и каналов связи. Также Desktop Grid характеризуются низкой стоимостью владения и большим вычислительным потенциалом. При этом внимания также заслуживают Desktop Grid корпоративного уровня — Enterprise Desktop Grid. Такие системы позволяют частным компаниям и научным организациям получить все преимущества использования высокопроизводительных вычислений без необходимости масштабных инвестиций в инфраструктуру, поддержку и сопровождение.

Список литературы

- [1] Foster I., Kesselman C. The Grid: Blueprint for a New Computing Infrastructure.— San Francisco: Morgan Kaufmann Publishers Inc., 1999.— 748 p.
- [2] BOINC: Программное обеспечение с открытым исходным кодом для организации добровольных распределённых вычислений и распределённых вычислений в сети. URL: <http://boinc.berkeley.edu/index.php> (дата обращения: 21.01.2014).
- [3] Distributed Computing Platforms. URL: <http://distributedcomputing.info/platforms.html> (дата обращения: 21.01.2014).
- [4] Project statistics — SETI@home. URL: <http://www.allprojectstats.com/po.php?projekt=15> (дата обращения: 21.01.2014).
- [5] Кратко о системе X-Com. URL: <http://x-com.parallel.ru/node/10> (дата обращения: 21.01.2014)
- [6] Thani D., Tannenbaum T., Livny M. *Distributed computing in practice: the Condor experience* // Concurrency and Computation: Practice and Experience, 2005. Vol. 17, p. 323–356.
- [7] Chien A., Calder B., Elbert S., Bhatia K. *Entropia: Architecture and performance of an enterprise desktop grid system.* // Parallel and Distributed Computing, 2003. Vol. 63, p. 597–610.
- [8] Luther A., Buyya R., Ranjan R., Venugopal S. *Alchemi: A .NET-Based Enterprise Grid Computing System* // Proceedings of the 6th International Conference on Internet Computing (ICOMP'05).— Las Vegas, USA, 2005.—p 43–53.
- [9] Chu X., Nadiminti K., Jin C., Venugopal S., Buyya R. *Aneka: Next-Generation Enterprise Grid Platform for e-Science and e-Business Applications* // Proceedings of the 3th IEEE International Conference on e-Science and Grid Computing (e-Science 2007).— Bangalore, India, 2007.—p. 151–159.
- [10] Choi S. J., Kim H. S., Byun E. J., Baik M. S., Kim S. S., Park C. Y., Hwang C. S. *Characterizing and Classifying Desktop Grid* // IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID 2007), Sixth International Workshop on Global and Peer to Peer Computing (GP2P).— Rio de Janeiro, Brazil, 2007.— p. 743–748.

- [11] Ивашко Е. Е., Головин А. С. *Методы Data Mining для анализа больших массивов данных в гетерогенной grid на базе BOINC* // Научный сервис в сети Интернет: поиск новых решений: Труды Международной суперкомпьютерной конференции (17-22 сентября 2012 г., г. Новороссийск).— М.: Издательство Московского государственного университета, 2012.— С. 31–34.
- [12] Renderfarm.fi, The Publicly Distributed Rendering Service. URL: <http://www.renderfarm.fi/> (дата обращения: 21.01.2014).
- [13] Weisbecker A., Pfreundt F., Linden J., Unger S. *Fraunhofer Enterprise Grids. Business Cases.*— Stuttgart: Fraunhofer IRB Verlag, 2008.— 62 p.
- [14] Ивашко Е. Е., Никитина Н. Н. *Использование BOINC-grid в вычислительном комплексе научных исследований* // Вестник Новосибирского государственного университета, 2013. Т. 11, №1, с. 53–57.

Рекомендовал к публикации

Программный комитет

Второго национального суперкомпьютерного форума НСКФ-2013

Об авторе:



Евгений Евгеньевич Ивашко

Кандидат физико-математических наук, научный сотрудник Института прикладных математических исследований Карельского научного центра РАН.

e-mail: ivashko@krc.karelia.ru

Образец ссылки на публикацию:

Е. Е. Ивашко. *Desktop Grid корпоративного уровня* // Программные системы: теория и приложения: электрон. научн. журн. 2014. Т. 5, № 1(19), с. 183–190.

URL: http://psta.psir.ru/read/psta2014_1_183-190.pdf

E. E. Ivashko. *Enterprise Desktop Grids*.

ABSTRACT. In this paper we consider Enterprise Desktop Grids, their capabilities, features, characteristics and applications. There are also examples of software aimed to construct Enterprise Desktop Grids.

Key Words and Phrases: Desktop Grid, Enterprise Desktop Grid, distributed computing.