

НАУЧНЫЕ СБОРЫ

Распределённые системы — велевание времени

30 ноября — 3 декабря в Институте вычислительных технологий СО РАН прошла XIII конференция «Распределённые информационные и вычислительные ресурсы» (DICR'2010).

Современные информационные технологии, позволяющие создавать, хранить, перерабатывать и обеспечивать эффективные способы представления информационных ресурсов потребителю, стали важным фактором жизни общества и средством повышения эффективности управления всеми сферами общественной деятельности. Стремительное развитие глобальных информационных и вычислительных сетей привело к изменению фундаментальной парадигмы обработки данных, направление которой переместилось на использование распределённых информационно-вычислительных ресурсов и поддержку инфраструктуры для свободного доступа к ним. В Институте вычислительных технологий СО РАН интерес к этим направлениям возник ещё в 1996 году. Тематика этих работ является одной из основных в деятельности научной школы ак. Ю.И. Шокина.

Нынешняя конференция, уже XIII по счёту, призвана стимулировать дальнейшее развитие моделей и методов обработки данных на основе современных информационных технологий. Актуальность данной тематики обусловливается ещё и всевозрастающей необходимостью проведения совместных работ группами территориально разобщённых учёных с возможностью обеспечения доступа к результатам исследований и распределённым базам данных. Обсуждение этих вопросов непосредственно связано с теоретическими и прикладными разработками в области информационных технологий, интеграции междисциплинарных данных и средств математического моделирования.

Основные направления работы конференции — организация распределённых вычислений и доступа к информационным ресурсам, в т.ч. на основе технологий Grid (среды для быстрого создания программного обеспечения с использованием глобально распределённых команд). На пленарных заседаниях и двух секциях были подробно рассмотрены технологии построения распределённых информационных систем, интеграции информационных ресурсов и их представления в удобном для пользователя виде, корпоративные интеллектуальные технологии обработки данных, геоинформационные системы для задач экологии и природопользования, электронные библиотеки и коллекции, библиометрические технологии и вебометрия, проблемы информационной безопасности, актуальные проблемы математического моделирования и обработки данных с использованием высокопроизводительных вычислительных систем.

На сайте конференции зарегистрировались 188 человек, было принято для публикации 111 докладов. В Институте вычислительных технологий собрались специалисты из научно-исследовательских учреждений и вузов Новосибирска, Томска, Москвы, Иркутска, Хабаровска, Красноярска, Кемерово, Самары, Барнаула, Омска, Владивостока, Южно-Сахалинска, а также из Белоруссии и Казахстана. Весьма внушительна научная квалификация участников: академик и 4 члена-корреспондента РАН,

41 доктор и 31 кандидат наук.

В рамках конференции прошёл тренинг, организованный московскими специалистами, представителями Международной федерации Grid-систем из персональных компьютеров (International Desktop Grid Federation, IDGF).

Основная концепция технологии Grid — это объединение всех ресурсов сети Интернет в единую интегрированную среду распределённых ресурсов, которая, возможно, составит информационно-вычислительную инфраструктуру будущего. Единый интерфейс должен предоставлять доступ ко всем необходимым ресурсам так, словно мы имеем дело с одним огромным «метакomпьютером».

Современные вычислительные Grid-системы можно разделить на два типа: сервисные и Grid из персональных компьютеров (Desktop Grid). Если взять современный домашний или офисный персональный компьютер и проследить время загрузки его процессора, то можно увидеть, что вычислительные мощности машины используются лишь на несколько процентов, а основная часть вычислительного ресурса простаивает. Технологии Desktop Grid позволяют использовать эти простаивающие ресурсы.

Происходит это следующим образом: владелец персонального компьютера устанавливает и настраивает на своей машине приложение, которое подключается к удалённому серверу, скачивает с него задания, запускает на своей платформе и возвращает результат обратно на сервер. Работа машины в рамках инфраструктуры Desktop Grid абсолютно не препятствует его работе в качестве персонального компьютера — задания запускаются с низким приоритетом и используют только незадействованное процессорное время. Данная технология используется для организации проектов добровольных вычислений. Кроме того, если отделить результат от приложения, то на основе этой технологии можно объединить имеющиеся персональные машины и задействовать их возможности.

Теперь несколько слов о молодой международной организации International Desktop Grid Federation (IDGF). Официально о её создании было объявлено в Амстердаме 15 сентября с.г. Федерация ставит целью всемерно способствовать объединению пользователей и разработчиков Grid-систем из добровольно выделенных персональных компьютеров (Volunteer Desktop Grids), распространяя информацию, занимаясь обучением, оказывая помощь в запуске новых проектов, и являясь открытой для всех желающих.

IDGF имеет интерактивный Web-портал и информационный центр с материалами на разных языках. На портале работает форум для общения на темы Desktop Grid. В настоящее время членами IDGF уже являются 20 организаций. Целями на ближайшие два года заявлены увеличение роли научных исследований, выполняемых с помощью Desktop Grid, и, по крайней мере, десятикратный рост числа компьютеров, используемых для таких вычислений.



Федерация проводит учебные курсы и распространяет обширный набор учебных материалов — от начальных руководств до описания технологий интеграции Desktop Grid в существующие распределённые вычислительные инфраструктуры. Тренинг в ИБТ — первый из обучающих семинаров IDGF, организованный в Сибири.

Во время занятий слушатели получили представление о Grid'e из персональных компьютеров и существующих реализациях данной технологии. Практическая часть занятий основывалась на технологии BOINC. Каждому участнику тренинга сначала предлагалось подключить тестовую машину к существующему специально созданному серверу добровольных вычислений и на практике увидеть, что представляет собой технология Desktop Grid с позиции добровольца, предоставляющего свой компьютер в качестве вычислительного ресурса.

На следующем этапе каждому из участников была предоставлена возможность попробовать себя в качестве администратора сервера и самостоятельно запустить распределённое приложение в рамках небольшой локально организованной вычислительной инфраструктуры. Для каждого участника тренинга был организован свой собственный проект вычислений. К нему можно было подключать вычислительные узлы и через систему электронного обучения Moodle (Modular Object-Oriented Dynamic Learning Environment) запускать распределённые приложения.

С просьбой поделиться своими впечатлениями о конференции и тренинге корреспондент «НВС» обратился к нескольким их участникам.

Александр Петрович Афанасьев, д.ф.-м.н., заведующий Центром Grid-технологий и распределённых вычислений Института системного анализа РАН, профессор МФТИ:

— Конференция оказалась очень работоспособной. Все ведущие институты за Уралом и, конечно, институты Академгородка представили очень серьёзные и крупные разработки. С содержательной точки зрения, мне кажется, конференция всех обогатила. Естественно, раз люди делают что-то интересное, возникает необходимость в общении. И я думаю, что количество контактов после этой конференции сильно увеличится.

Мы представляли здесь две группы наших разработок. Первая — это вычислительная среда для математиков, которая называется Mathcloud — «математическое облако». Идея эта разработка в некотором смысле переключается с модными сейчас облачными вычислениями. Это высокоуровневая среда с доступом через Интернет, которая представляет разработчикам возможность решать не очень сложными средствами сложные вычислительные задачи. Понятно, что хороший физик далеко не обязательно должен быть хорошим математиком, хорошо вычислителем. Мы пытаемся дать возможность этим специалистам достаточно простыми средствами использовать сложный математический аппарат, применять современные вычислительные алгоритмы, которые могут распараллеливаться,

распределяться и вообще могут быть задействованы на очень высокопроизводительной среде.

Вторая тема, с которой мы выступаем — это тренинг по Desktop Grid'y. Это довольно старая идея, которая развивается уже давно — идея объединить рабочие станции в единую вычислительную среду. Но в настоящее время в Евросоюзе возникла инициатива создания не просто Grid'a из рабочих станций, а создание сети, ориентированной на научные исследования. Инициатива заключается в том, что именно научные организации, наиболее подготовленные, наиболее компетентные, объединяют свои рабочие станции в эту единую среду. Не надо понимать, что это только десктопы, простые рабочие компьютеры. Сюда можно включить и небольшие кластеры, которых сейчас много везде, устройства и вычислительные установки, основанные на графических ускорителях. Т.е. среда позволяет объединять достаточно разнородные ресурсы.

Перспектива заключается в следующем. Понятно, что сегодня вычислительная техника развивается очень быстро. Мы уже имеем персональные многоядерные компьютеры. Через три-четыре года у каждого из нас на столе будет стоять уже небольшой суперкомпьютер. Одновременно растёт скорость телекоммуникационной среды — общение между компьютерами становится очень быстрым. Поэтому, если всё это объединить по соответствующим протоколам, что мы сейчас и обсуждаем с коллегами, получается фантастической мощности вычислительная среда, вполне конкурентоспособная с самыми мощными суперкомпьютерами.

Очень хотелось бы активно начать присоединяться к этой европейской инициативе. В этом смысле мы очень сильно рассчитываем на Академгородок, поскольку это компактная высокопрофессиональная среда, достаточно хорошо оснащённая и вполне обозримая. Ну и народ здесь такой, который быстро реагирует. Так что мы надеемся, что получим здесь много союзников.

Desktop Grid даёт колоссальные вычислительные преимущества и избавляет нас от тех хлопот, которые сейчас многие получают в связи с установкой многофлпных суперкомпьютеров. Известны примеры, что где-то устанавливается очень мощный суперкомпьютер, а энергоснабжение недостаточно, и компьютер молчит. Он молчит год-два-три и за это время морально устаревает. И вообще все компьютеры морально устаревают. Но парк персональных компьютеров обновляется автоматически! За этим не нужно следить специально — люди сами меняют свои компьютеры. Поэтому Desktop Grid — очень перспективная вещь. Единственный конкурент этому, как сейчас говорят — это развитие идеологии облачных вычислений, когда вообще не нужен компьютер, а я могу зайти в виртуальную среду, получить там необходимую мне вычислительную мощность или необходимый сервис, который обработает любую задачу. Может быть, и так произойдёт. Но, мне кажется, по крайней мере, в ближайшие 10 лет эти два проекта будут сосуществовать параллельно.

