

М. В. Колпачков, Н. А. Петрин – студенты кафедры компьютерной математики и программирования
М. Д. Поляк – научный руководитель

РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛЕНИЯ

В наше время тема распределенных и облачных вычислений актуальна как никогда. В данной работе мы рассмотрим, чем характеризуются оба типа систем, их особенности, области применения в мировой практике.

Распределённые вычисления

Понятие GRID (англ. *grid* — решётка, сеть) подразумевает использование множества вычислителей, объединенных в единую сеть и решающих поставленную задачу сообща. Особенности таких систем в отличие от классических являются:

- разбиение вычислений на подзадачи;
- высокая скорость в сложных расчетах;
- масштабируемость;
- отказоустойчивость.

Остановимся на каждом пункте более подробно.

Применяя разбиение вычислений на подзадачи, мы распараллеливаем решение нашей задачи, тем самым повышая скорость вычислений. Такие системы особенно эффективны при решении NP-полных задач, обработки изображений, например, ДЗЗ, речевых сигналов, статистическом анализе, машинном обучении (нейронные сети), генетических алгоритмах, биоинформатике и т.д.

С приходом широкополосного интернета появилась возможность отдавать часть ресурсов своего ПК на нужды научных проектов. В качестве примера можно привести проект Folding@Home, который на 10 ноября 2011г. имеет мощность 6 петафлопс. Для сравнения, первую строчку в мировом рейтинге суперкомпьютеров TOP500 занимает система «K computer» с мощностью около 10,5 петафлопс, второе место у «Tianhe-1A» (2,5 петафлопс).

Такие системы могут быть легко масштабируемы, т.к. для введения нового узла в сеть необходима лишь единица вычислительной техники, с установленным специальным ПО. Добавление новых узлов происходит с минимум предварительных настроек. Ограничения же на масштабируемость, вернее невозможность увеличения быстродействия вычислений для конкретной задачи, ограничена Законом Амдала, который гласит, что суммарное время выполнения задачи, разделенной на несколько частей, на параллельной системе не может быть меньше времени выполнения ее самого длинного фрагмента, т.е.

$$S \leq \frac{1}{f + (1 - f) / p},$$

где S – получаемое ускорение, p – количество процессоров, f – доля операций, выполняющихся последовательно (при этом под долей понимается время выполнения последовательной программы).

Отказоустойчивость, в зависимости от архитектуры, обеспечивается либо резервированием, либо специальными управляющими серверами, которые управляют контролем выполнения заданий.

Также хотелось бы отметить, что на сегодняшний день производительность мобильных устройств значительно возросла. Мы предлагаем использовать их вычислительные ресурсы в ночное время, во время зарядки и при включенном Wi-Fi соединении.

Достоинства:

- технология Grid систем позволяет строить мощные вычислительные комплексы без относительно больших финансовых вложений по сравнению с кластерными системами;
- использование распределенных вычислений позволяет добиться высокой отказоустойчивости;

- возможность эффективной обработки больших объемов данных за счет уменьшения накладных расходов на перемещение данных к вычислителям.

Недостатки:

- сложность применения технологии для решения задач реального времени;
- необходимость перепроектирования ПО для распределения вычислений.

Облачные вычисления

Облачные вычисления (англ. cloud computing) – модель устройства вычислительной системы, при которой, клиент, не обладая значительными вычислительными ресурсами (необходимыми объемами памяти и т.д.), получает возможность использовать таковые на сервере (облаке).

В такой системе клиент выступает руководствующим элементом системы, управляя запуском, и прекращением работы приложения, которое можно запускать, словно на локальной машине.

Облачные системы отличают следующие характеристики:

- объединение вычислительных ресурсов;
- универсальный доступ к облаку;
- учёт использования вычислительных мощностей.

Объединение вычислительных ресурсов в таких системах обозначает единый интерфейс для доступа ко всем ресурсам системы, при этом нагрузка автоматически распределяется между узлами, а клиент может управлять, по сути, только желательным географическим положением узла (для повышения скорости обмена данными).

Универсальность доступа к облаку заключается в том, что к такой системе можно получить доступ и с ПК, и с мобильного устройства, например, современного телефона.

Учёт использования системы клиентом происходит автоматически на некотором уровне абстракции (объём данных, количество транзакций и т.д.) и на основе этих данных оценивается задолженность клиента.

Облачные системы можно классифицировать по типу облака:

- частное облако;
 - публичное облако;
 - гибридное облако
- и по методу обслуживания:

- аренда ПО (SaaS);
- аренда исполнительной среды (PaaS);
- аренда вычислительной инфраструктуры (IaaS).

Частные облака предназначены для использования одной организацией, возможно клиентами и подрядчиками организации. Такое облако может находиться в собственности как самой организации, так и третьих лиц.

Публичные облака предназначены для использования широкой публикой. Такое облако находится во владении крупных коммерческих, научных и правительственных организаций. Физически такое облако существует в юрисдикции владельца – поставщика услуг.

Гибридное облако – облако из нескольких облаков (частных, публичных), которые остаются уникальными, но связаны. Примером такого облака может служить облако из частных облаков, объединённых для балансировки нагрузки между собой.

Облачные системы аренды ПО – такие системы, которые предоставляют клиенту доступ к какому-то определённому прикладному ПО, которое он (клиент) может запускать с любых устройств за счёт унифицированного интерфейса (веб-страница) или клиента для его платформы (под платформой можно понимать как ОС, так и различные типы вычислительных устройств).

Облачные системы исполнительной среды – системы, которые предоставляют клиенту нечто вроде унифицированной вычислительной машины с готовыми инструментальными средствами создания, тестирования и отладки прикладного ПО, с возможностью развёртывания на нём нового или уже

существующего ПО. В такой системе весь низкоуровневый контроль над распределением вычислительных мощностей осуществляется провайдером, предоставившим сеть.

Облачные системы аренды вычислительной инфраструктуры – такие системы, которые предоставляют клиенту унифицированный интерфейс к системе и дают ему развёртывать любое ПО, вплоть до операционных систем. Клиент, таким образом, получает полный контроль над этими виртуальными системами, не имея при этом контроля над физическими составляющими системы.

Сегодня Облачные системы активно применяются владельцами мобильных устройств, которые не отличаются высокой производительностью и большими объёмами памяти. Облачные системы позволяют, заплатив один раз провайдеру, иметь доступ с мобильного телефона к любому “тяжёлому” ПО, например, Matlab, или современным компьютерным играм. Кроме того, имеется возможность хранить объём данных сильно больший, чем доступный на мобильном устройстве.

Примерами облачных систем могут служить многие современные проекты, выделим следующие:

- Amazon EC2 – облачный сервис, предоставляющий простой интерфейс для выбора, настройки и эксплуатации ПО клиентом. Предоставляет полный контроль над вычислительными ресурсами вместе с доступной средой работы;

- DropBox – облачный сервис распределённого хранения данных в интернете.

Достоинства:

- облачные системы предоставляют обобщенный интерфейс для решения широкого круга задач;

- использование облачных вычислений позволяет добиться высокой отказоустойчивости.

Недостаток – сложность применения технологии для решения задач реального времени.