

Гибридная суперкомпьютерная система*

Р.Т. Файзуллин, А.А. Свенч, В.А. Соловьев, В.Ф. Феллов, И.Г. Хныкин

Омский государственный технический университет

В статье рассматривается гибридная суперкомпьютерная система на базе кластера центральных процессоров (ЦПУ), кластера графических процессоров (ГПУ) и системы хранения данных (СХД). Описывается программно-аппаратная база, используемая для построения. Формируется класс задач, которые могут оптимально выполняться с использованием гибридной системы. Рассматриваются прикладные задачи криптографического анализа, моделирования транспортных потоков в сложных системах, исследования адсорбции сложных органических молекул.

1. Введение

В настоящее время есть множество общедоступных вычислительных кластеров (суперкомпьютеров). Одни основаны на стандартной архитектуре CPU, другие на менее известной архитектуре GPU. Проведенный обзор показал, что вычислительных систем, объединяющих данные архитектуры в единый кластер, нет, либо они являются коммерческим секретом.

Архитектура CPU позволяет использовать практически любые шаблоны параллельных вычислений, в то время как GPU является менее гибким, но гораздо более производительным вариантом для решения многих вычислительных задач. В частности, архитектура современных GPU NVidia Tesla включает в себя множество масштабируемых блоков, не обладающих мощной управляющей логикой и большим объемом кэш-памяти. Эта архитектура может эффективно применяться при вычислениях с большим параллелизмом и интенсивной арифметикой. Все функции, выполнимые на GPU, не поддерживают рекурсии и имеют некоторые другие ограничения, которых нет в архитектуре CPU. С помощью созданной авторами модели программирования стало возможным разрабатывать сложные проекты с использованием параллельных вычислений, которые одновременно могут использовать «гибкость» CPU и более скоростные вычисления GPU.

Другим недостатком существующих суперкомпьютерных систем является непрозрачный доступ пользователя к своим проектам на кластере. Сегодня пользователю предлагается обучиться работать с арсеналом программного обеспечения (программой удаленного доступа, командной строкой, компилятором, планировщиком и т.д.), установленным на целевом кластере. Поэтому авторами разработана программная система, позволяющая осуществлять удаленную работу с проектами без необходимости погружаться в устройство операционной системы кластера.

2. Описание технологии и аппаратных средств

Гибридная суперкомпьютерная система включает в себя:

– Вычислительную систему, объединяющую кластер CPU(центральных процессоров) x86/64 – на базе процессоров Intel Xeon, и кластер GPU(графических процессоров) – на базе процессоров NVidia Tesla.

– Модуль системы хранения данных на базе оборудования Sun Microsystems.

– Инновационную модель программирования, позволяющую объединить вычисления, использующие CPU и GPU.

– Инновационную программную систему, которая управляет выполнением проекта, использующего гибридные вычисления, и предоставляет прозрачный интерфейс для работы пользователей с системой на базе web-технологий.

* Работа поддержана грантом конкурса «У.М.Н.И.К.», конференция «ОМСКОЕ ВРЕМЯ – ВЗГЛЯД В БУДУЩЕЕ», 2010 г.

На сегодняшний день в распоряжении ОмГТУ находится пять вычислительных узлов с архитектурой CPU (mgr, cn01, cn02, cn03, cn04). Каждый вычислительный узел включает два 4-ядерных процессора HP X5472 DL 160G5. Пиковая производительность достигает 1 Tflop, а объем оперативной памяти составляет 40GB. Структура масштабируема и расширяема. Вычислительная мощность может быть увеличена за счет добавления дополнительных узлов.

В рамках гибридизации к суперкомпьютеру добавлен вычислительный узел cn05, осуществляющий управление кластером GPU. Для его построения выбрана платформа NVidia Tesla 10 с архитектурой CUDA GPU. Узел представляет собой персональный компьютер с подключенными к нему ячейками NVidia Tesla S1070, каждая из которых включает в себя 4 GPU с суммарной пиковой вычислительной мощностью 4 Tflops в операциях с одинарной точностью. То есть производительность всей системы, при относительно малых затратах увеличивается до 5 Tflops (при задействовании одной ячейки Tesla S1070). Для выполнения параллельных вычислений используется NVidia CUDA API для языков программирования C, C++, Fortran [2,3].

Другим дополнением суперкомпьютерной системы является модуль хранения данных (ssd) на базе системы Sun StorageTek 9900V. Данная система отличается высокой отказоустойчивостью.

Все узлы объединены в локальную высокоскоростную сеть (1Gb/сек). Параллельное выполнение программ осуществляется с помощью технологий MPI и CUDA. Поддерживаемые реализации MPI: OpenMPI, HPMPI, MPICH [1].

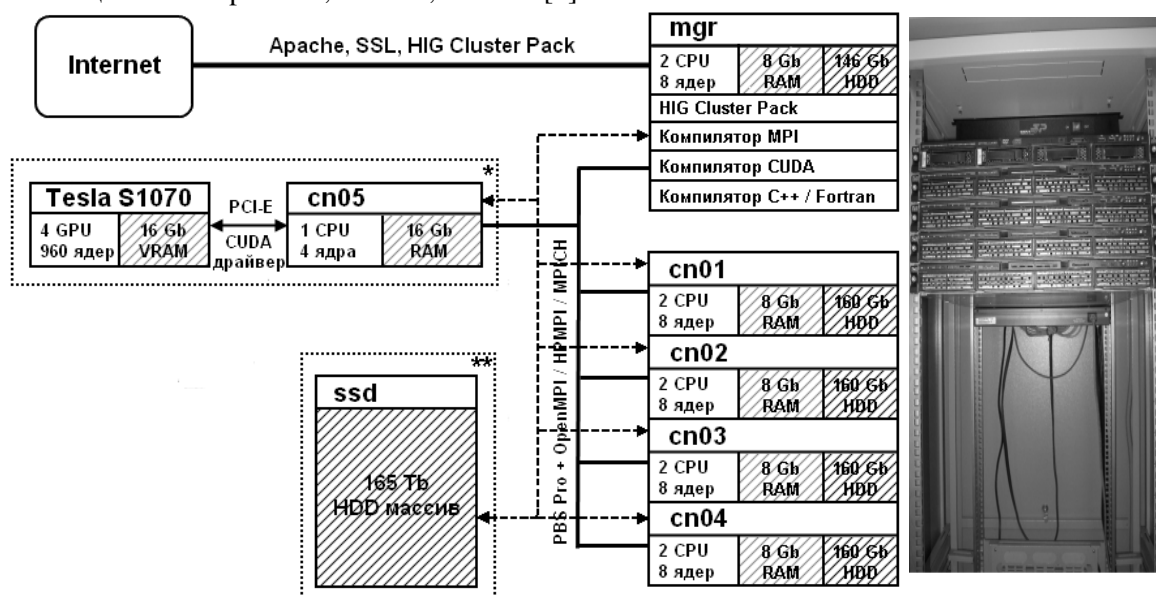


Рис.1. Схема суперкомпьютерной системы ОмГТУ.¹²

Для управления кластером один из его узлов выделен как управляющий (mgr), при необходимости он может быть использован для вычислений. На управляющем узле установлена авторская программная система для управления проектами на кластере (HIG_Cluster_Pack). С её помощью узел принимает и выполняет команды авторизованных пользователей через Интернет. Все соединения осуществляются по защищенному HTTPS протоколу (см. Рис.1).

¹ На момент написания статьи узел cn05 смоделирован с помощью персонального компьютера с 4 установленными графическими адаптерами NVidia GeForce GTX260.

² На момент написания статьи дисковый массив ssd имеет общий объем 12 Tb, планируется расширение до 165 Tb.

3. Программная система прозрачного доступа к кластеру

Программная система доступа к кластеру разработана авторами проекта и основана на базе web-сервера Apache. Данная система является инновационной полностью переносимой и расширяемой и может работать на вычислительных кластерах с любой аппаратной конфигурацией, операционной системой и программным обеспечением. На сегодняшний день данная система (HIG_Cluster_Pack) не имеет аналогов и успешно используется на кластере ОмГТУ.

Программа предоставляет следующие возможности:

– Трехфакторная система авторизации на уровне web-сервера, самой программы и операционной системы.

– Для каждого пользователя создается, так называемая, «песочница». То есть, все исполняемые файлы проекта могут обращаться только к ресурсам своего проекта и таким образом исчезает возможность злонамеренного повреждения кластера или проектов других пользователей.

– Пользователю предоставляется удобный web-интерфейс, который позволяет создать/удалить проект, загрузить проект, скомпилировать проект, запустить проект, посмотреть файлы проекта и т.п. При этом поддерживаются все необходимые настройки работы с проектом. Пользователю нет необходимости разбираться в опциях компилятора и планировщика заданий.

– Существует система полуавтоматической обработки заданий, когда пользователь с помощью определенных HTTPS запросов управляет работой проекта на кластере. Таким образом, появляется возможность создания программ, работающих на персональном компьютере и использующих вычислительные мощности кластера для просчета ресурсоемких блоков.

– Благодаря дружественному интерфейсу регистрация нового пользователя производится удобным как самому пользователю, так и администратору кластера способом. Пользователю достаточно сгенерировать запрос на сертификат доступа, где указываются необходимые данные. Администратору достаточно разрешить работу на кластере с данным сертификатом. Все остальные действия выполняются автоматически.

4. Прикладные задачи

Суперкомпьютерная система, смоделированная авторами в виде рабочего прототипа, ориентирована на широкий круг задач. В сочетании с инновационными авторскими разработками по созданию унифицированных средств доступа и управления, данная система позволяет использовать сразу все ведущие решения в области высокопроизводительных вычислений.

4.1. Математическое моделирование транспортных потоков на основе микроскопической схемы предиктор-корректор

Представляется возможным использование параллельных вычислений в задаче моделирования транспортных потоков, которая отличается большой сложностью в случае моделирования и оптимизации работы транспортной сети крупного города. Существующие модели, макроскопические и микроскопические в настоящее время ограничены однопроцессорными реализациями.

В качестве топографической основы модели транспортной системы города рассмотрим неориентированный граф, ребра это дороги или магистрали, узлы это перекрестки. Будем считать, что движение везде двухполосное, разделенное сплошной линией, т.е. обгоны запрещены и на всех перекрестках стоят светофоры. Транспортные средства – это точки, расположенные на ребрах. В начальном состоянии системы все транспортные средства имеют нулевую скорость. Для каждого транспортного средства случайно выбираются пункты назначения – точки на каком-либо ребре, для которых вычисляется оптимальный маршрут. В

результате поиска пути, каждому транспортному средству с номером i_s , поставлен в соответствие массив номеров ребер $ir_s(l)$, которые он должен пройти. Здесь $l = 1, \dots, L$, где L это число ребер в графе. Время считаем изменяющимся дискретно и вводим два характерных величины для времени Δt и $\Delta \tau$, где первая величина намного больше второй. Мы предполагаем, что Δt , это общий интервал времени, на которое независимо прогнозируется движение каждым водителем, а $\Delta \tau$, это шаг по времени после которого водитель вынужден корректировать свой прогноз.

Решение частных задач моделирования и оптимизации естественным образом приводит нас к требованию максимального повышения производительности вычислений. Как показывают расчеты, движение в режиме реального времени транспортных средств в количестве 100 000 единиц моделируется одним процессором, но задачи моделирования большей размерности и особенно оптимизационные задачи требуют повышения скорости вычислений на порядки.

Например, если речь идет об управлении движением транспорта, то число расчетов по необходимости должно быть велико, и они должны происходить существенно быстрее.

Как показывает опыт вычислений, число контролируемых транспортных средств, при подобном моделировании может достигать чисел порядка 10^6 и время расчета вариантов движения существенно меньше, чем время реализаций этих вариантов в действительности. Также, представляется перспективным перенести часть массовых вычислений на графические процессоры, выбирая оптимальные размеры региона или специальным образом распараллеливать вычисления на обычные и графические процессоры [4].

4.2. Исследование многоцентровой адсорбции молекул с возможностью различной ориентации в адсорбционном слое

В последнее время наблюдается значительный рост исследований, направленных на изучение поведения молекул, адсорбированных на поверхности твердых тел. С одной стороны это во многом обуславливается значительным прогрессом в техническом обеспечении экспериментов (сканирующая туннельная микроскопия, атомно-силовая микроскопия и т.д.). С другой стороны необходимостью в переходе на качественно другой уровень в создании микро(нано)электронных приборов (полевых транзисторов, органических светодиодов, нелинейной оптики, газовых сенсоров, хранителей информации, микропроцессоров...) – на уровень атомов, молекул и их ансамблей. Данные экспериментов показывают, что при адсорбции молекулы занимают несколько активных центров поверхности и вследствие своей сложной структуры могут по-разному ориентироваться в адсорбционном монослое [5]. Несмотря на это, теоретических работ по исследованию подобных систем в свете многоцентровой адсорбции с возможностью различной ориентации молекул на поверхности очень мало.

Для построения модели адсорбционной системы была использована МРГ (модель решеточного газа). В рамках МРГ поверхность представляла собой двумерную решетку с квадратной симметрией. В качестве модельной молекулы-адсорбата был выбран димер – молекула, состоящая из двух одинаковых сегментов, расстояние между которыми равно постоянной решетки. Димер не обладает сферической симметрией и при адсорбции может занимать один или два активных центра поверхности (перпендикулярно или параллельно поверхности). Моделирование проводилось в большом каноническом ансамбле. Параметрами модели являлись: химический потенциал, температура, разница между теплотами адсорбции димера на два и на один активный центр, размер решетки. Модель исследовалась при помощи методов Монте-Карло и трансфер-матрицы.

Авторами модели были построены графики функции степени покрытия поверхности и изотермы адсорбции, на которых видны два ярко выраженных плато. Этим плато соответствуют упорядоченные фазы, возникающие в системе. В области низкого значения химического потенциала на поверхности наблюдается упорядоченная структура, состоящая только из димеров, адсорбированных на два активных центра. В области высокого значения химического потенциала – только из димеров, адсорбированных на один активный центр.

Интересно, что структур, состоящих одновременно из параллельно и перпендикулярно ориентированных димеров, на поверхности не образовывалось. Также были построены графики дифференциальной теплоты адсорбции и внутренней энергии системы от плотности монослоя, из которых видно, что в системе наблюдаются два режим адсорбции.

Для каждой упорядоченной структуры был введен свой уникальный параметр порядка, который равнялся единице, если в системе имела место соответствующая упорядоченная фаза, и стремился к нулю, если ее не было. Исходя из значений параметров порядка, была построена фазовая диаграмма.

Расчет модели, в силу большой вычислительной емкости и хорошей распараллеливаемости лежащих в ее основе численных методов (в частности, метода Монте-Карло) производился с помощью кластера ЦПУ, являющегося подсистемой гибридной суперкомпьютерной системы.

Литература

1. Стандарт MPI-2.0 // <http://www.mpi-forum.org/docs/docs.html>
2. CUDA Reference manual // <http://developer.download.nvidia.com>
3. NVidia Developer zone // <http://developer.nvidia.com>
4. Файзуллин Р.Т., Свенч А.А., Хныкин И.Г. Применение гибридной суперкомпьютерной системы в задачах криптоанализа // Доклады ТУСУР / Томск: ТУСУР, №1 (21), часть 1. 2010. С. 61-63.
5. Su G. J., Zhang H. M., Wan L. J., Bai C. L., Wandlowski T. Potential-induced phase transition of trimesic acid adlayer on Au(111) // J. Phys. Chem. B. – 2004. – V.108. – P.1931 – 1937.