

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС НА БАЗЕ ВЫЧИСЛИТЕЛЬНЫХ КЛАСТЕРОВ РАЗЛИЧНЫХ АРХИТЕКТУР

А.А. Свенч, Р.Т. Файзуллин

Распространенные на данный момент высокопроизводительные вычислительные системы специализированы и ориентированы на решение определенного класса задач, как то: инженерно-технических задач, задач логистики, задач обработки изображений и т.д. Такие системы базируются либо на вычислительных узлах CPU (центральных процессорах), либо на узлах GPU (графических процессорах).

В данной работе рассматривается программно-аппаратный комплекс - суперкомпьютерная система, включающая оборудование различных производителей (вендоров) для эффективного решения широкого класса задач, связанных с оптимизацией производства и бизнес процессов. 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.

Описание аппаратной реализации

В качестве подсистемы CPU используются вычислительные узлы HP X5472 DL 160G5. Каждый вычислительный узел включает два 4-ядерных процессора Intel Xeon. Пиковая производительность одного узла достигает 200 GFlop (млрд. операций с плавающей точкой в секунду), а объем оперативной памяти составляет 8GB. Структура масштабируема и расширяема. Вычислительная мощность может быть увеличена за счет добавления дополнительных узлов.

Для построения подсистемы GPU выбрана платформа NVidia Tesla 10 с архитектурой CUDA. 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.

Все узлы (CPU, GPU, система хранения данных) объединены в локальную высокоскоростную сеть (1Gb/сек).

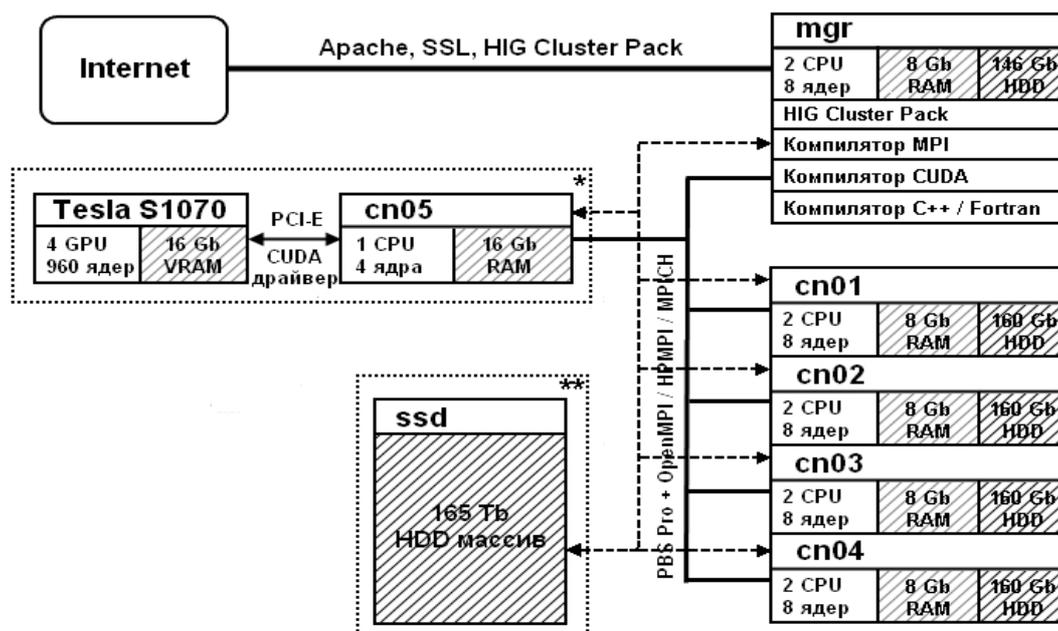


Рис. 1. Схема прототипа суперкомпьютерной системы.

Для управления суперкомпьютерной системой один из узлов CPU выделен как управляющий (mgr), при необходимости он может быть использован для вычислений. На управляющем узле установлена авторская программная система для управления проектами (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).

Шаблон объединенного параллельного программирования

Разработанный авторами шаблон параллельного программирования позволяет использовать распределенные вычисления на графических процессорах GPU внутри потоков классической параллельной программы. Рассмотрим параллельную программу, использующая технологию MPI. Пусть в одном или нескольких потоках данной программы имеются вычислительные блоки, которые можно распараллелить с помощью технологии CUDA. Пользователю необходимо выделить данные блоки и с помощью специальных инструкций CUDA (расширения языка программирования, предоставляемого производителями графических процессоров) запрограммировать их для вычислений на GPU. Остальные действия за пользователя сделает компилятор. Настройка компилятора сама выделит данные блоки в исходных текстах программы и вставит необходимые инструкции в бинарный код. Когда поток параллельной программы переходит к выполнению инструкций GPU, он перенаправляется на вычислительный узел кластера с GPU, где и производятся вычисления. Далее поток программы продолжает выполняться в обычном режиме.

Программная система прозрачного доступа к кластеру

Программная система доступа к кластеру разработана авторами проекта и основана на базе web-сервера Apache. Данная система является 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. Например, (с определенными оговорками) существует возможность использовать практически любой язык программирования при написании проектов. Пользователю нет необходимости разбираться в опциях компилятора и планировщика заданий. Грубо говоря, достаточно нажать две кнопки: «Скомпилировать» и «Запустить».
- 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 обычном персональном компьютере и использующих вычислительные мощности кластера для просчета ресурсоемких блоков посредством сети Интернет.
- 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.