



NVIDIA

**CUDA[®] АЛЬМАНАХ
СЕНТЯБРЬ 2014**



СОДЕРЖАНИЕ

НОВОСТИ NVIDIA CUDA

Конференция и тренинг по технологиям программирования графических процессоров
22-24 октября 2014 года в ИПМ им. Келдыша РАН [3](#)

Встреча GPU Computing Moscow 8 октября в Центре робототехники МГТУ им. Баумана [4](#)

Библиотека cuDNN упрощает ускорение приложений глубокого обучения [5](#)

Суперкомпьютерный консорциум университетов России и Российская Академия Наук
готовят к изданию книгу [6](#)

Отчет о производительности CUDA 6.5 [7](#)

Стань частью GTC 2015 [8](#)

ВЕБИНАРЫ НА АНГЛИЙСКОМ ЯЗЫКЕ [9](#)

НАУЧНЫЕ РАБОТЫ С ИСПОЛЬЗОВАНИЕМ ВЫЧИСЛЕНИЙ НА CUDA

Модельно-ориентированный подход к сквозному проектированию адаптивных распределенных сенсорных
сетей медицинского назначения // Володин К.И., Переходов А.И. [10](#)

Использование ГПУ процессоров для классификации запросов
при DDoS-атаке // А.И. Рейбандт, С.А. Сартин [11](#)

Параллельная версия жадного алгоритма кластеризации // М.А. Баранов [12](#)

Обзор методов упрощения полигональных моделей на графическом процессоре // А. В. Хоперсков,
С. А. Хоперсков, С. С. Храпов, А. В. Мишустин, А. В. Моисеев [12](#)

ПОЛЕЗНЫЕ РЕСУРСЫ ПО CUDA [13](#)

ВАКАНСИИ CUDA [14](#)

КОНТАКТЫ [16](#)

КОНФЕРЕНЦИЯ И ТРЕНИНГ ПО ТЕХНОЛОГИЯМ ПРОГРАММИРОВАНИЯ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ 22-24 ОКТЯБРЯ 2014 ГОДА В ИПМ ИМ. КЕЛДЫША РАН

Приглашаем разработчиков ПО, научных сотрудников, преподавателей, аспирантов и студентов принять участие в конференции «Дни технологий NVIDIA» и тренинге по технологиям программирования графических процессоров. Участники смогут ознакомиться с технологиями для высокопроизводительных вычислений CUDA и OpenACC, получить навыки программирования GPU, освоить продвинутые методики оптимизации GPU-кода, также в рамках конференции будет представлена платформа для решения вычислительных задач во встраиваемых и мобильных системах, автопилотах и роботах, будет представлена технология для организации виртуальных рабочих мест с полноценной графикой

Место проведения: Москва, Миусская пл., 4 (ИПМ им. М.В.Келдыша РАН).

Даты проведения конференции-тренинга: 22-24 октября 2014.

Запланированы выступления представителей NVIDIA, ИПМ им. М.В.Келдыша РАН, курс по программированию GPU и мастер-класс по оптимизации кода для GPU.

Участие бесплатное.

РЕГИСТРАЦИЯ ОБЯЗАТЕЛЬНА



ВСТРЕЧА GPU COMPUTING MOSCOW 8 ОКТЯБРЯ В ЦЕНТРЕ РОБОТОТЕХНИКИ МГТУ ИМ. БАУМАНА

Присоединяйтесь к группе GPU COMPUTING MOSCOW 8 октября для обсуждения тем, связанных с массивно-параллельными вычислениями, HPC, программированием GPU и применением вычислений на GPU в различных прикладных и научно-исследовательских сферах!

Программа:

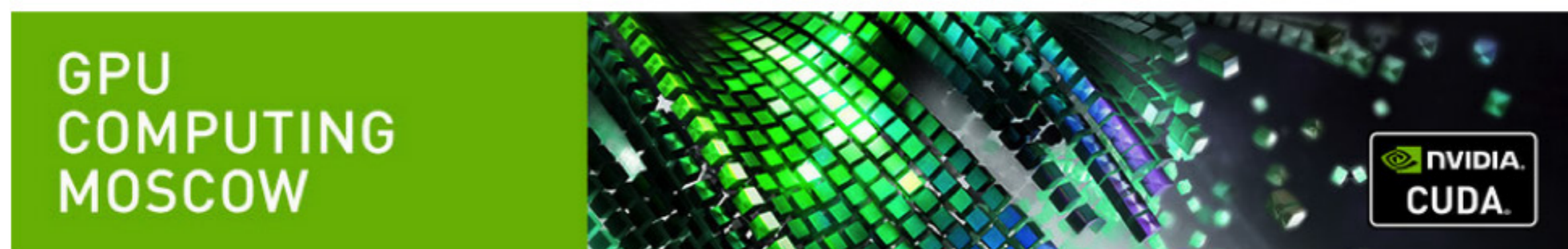
CUDA 6.5

Первый в мире встраиваемый суперкомпьютер Jetson TK1

Доклады участников

Место: Встреча пройдет в Центре робототехники МГТУ им. Баумана (Измайловская площадь, д.7). Начало в 18:00

[РЕГИСТРАЦИЯ ОБЯЗАТЕЛЬНА](#)



БИБЛИОТЕКА CUDNN УПРОЩАЕТ УСКОРЕНИЕ ПРИЛОЖЕНИЙ ГЛУБОКОГО ОБУЧЕНИЯ

NVIDIA представила простое в развертывании программное обеспечение, призванное помочь разработчикам направить мощь графических процессоров на ускорение работы приложений глубокого обучения (Deep Learning) в таких областях, как классификация изображений, видеоаналитика, распознавание речи и обработка естественного языка.



NVIDIA® cuDNN, мощная библиотека программирования на базе модели параллельного программирования CUDA®, ускоряет процессы глубокого обучения до 10 раз с помощью графических процессоров по сравнению с методами, выполняющимися только на CPU. Благодаря простоте использования, cuDNN позволяет разработчикам быстро создавать и оптимизировать новые модели и разрабатывать приложения более высокой точности, используя вычислительный потенциал GPU-ускорителей.

Глубокое обучение – это быстро развивающаяся область машинного обучения, которая включает создание сложных, многоуровневых, «глубоких» нейронных сетей. С помощью таких сетей мощные компьютеры учатся распознавать паттерны и объекты через анализ большого объема данных для обучения.

Графические процессоры все чаще применяются для ускорения приложений глубокого обучения, по мере того как исследователи и программисты осознают огромные преимущества GPU в ускорении процессов тренировки с большими объемами данных.

Исследователи из Калифорнийского Университета в Беркли интегрировали cuDNN в Caffe – одну из самых распространенных в мире инфраструктур для создания приложений глубокого обучения.

Кроме того, более 90% команд и трое из четырех победителей престижного конкурса 2014 ImageNet Large Scale Visual Recognition Challenge также использовали графические процессоры для ускорения своих проектов по глубокому обучению.

Подробнее об интеграции cuDNN в инфраструктуру глубокого обучения Caffe смотрите [на сайте cuDNN](#).

Подробнее о преимуществах GPU-ускорения для машинного обучения смотрите [на сайте NVIDIA](#).

Подробнее о событии – [в нашем блоге](#).

СУПЕРКОМПЬЮТЕРНЫЙ КОНСОРЦИУМ УНИВЕРСИТЕТОВ РОССИИ И РОССИЙСКАЯ АКАДЕМИЯ НАУК ГОТОВЯТ К ИЗДАНИЮ КНИГУ

Что это? Яркое и доступное описание примеров использования суперкомпьютерных технологий для решения самого широкого спектра задач промышленности, науки и образования. Жанр книги – подборка небольших статей и рассказов, изложенных в научно-популярной форме, о реальном применении передовых суперкомпьютерных технологий в различных областях.

Какова цель книги? Показать, что:

- суперкомпьютерные технологии не экзотика, они давно и активно используются во многих областях;
- суперкомпьютерные технологии и компьютерное моделирование являются необходимым инструментом для науки, промышленности, образования, для общества в целом;
- суперкомпьютерные технологии используются для получения качественно нового знания, результата, продукта;
- суперкомпьютерные технологии носят междисциплинарный и универсальный характер;
- суперкомпьютерные технологии мотивируют развитие науки, дополняя теорию и экспериментальные методы новыми вычислительными подходами к исследованиям: создание новых методов вычислительной математики, разработка новых подходов к технологиям программирования, новые идеи построения и принципы конструирования компьютерных систем, новые подходы к решению задач;
- мир суперкомпьютерных технологий разнообразен: архитектуры, компьютеры, технологии программирования, приложения на практике, размеры и размерности задач.
- это компьютерный мир, который существует параллельно с привычным компьютерным миром, закладывая основу его развития;
- суперкомпьютерные технологии – это самостоятельная область IT-мира, которая имеет особую ценность и развивается по своим законам;
- суперкомпьютерные технологии принципиально важны для развития высокотехнологичного общества и должны развиваться по отдельным программам;
- развитие суперкомпьютерных технологий – это забота государства, как транспортная инфраструктура, оборона, образование, освоение космоса...

Немаловажная цель:

- привлечь общее внимание, в частности, молодежь к новому, яркому, необычному и перспективному суперкомпьютерному направлению в теории и на практике.

Узнать подробнее и принять участие: http://hpc-russia.ru/book_superproblems6.html

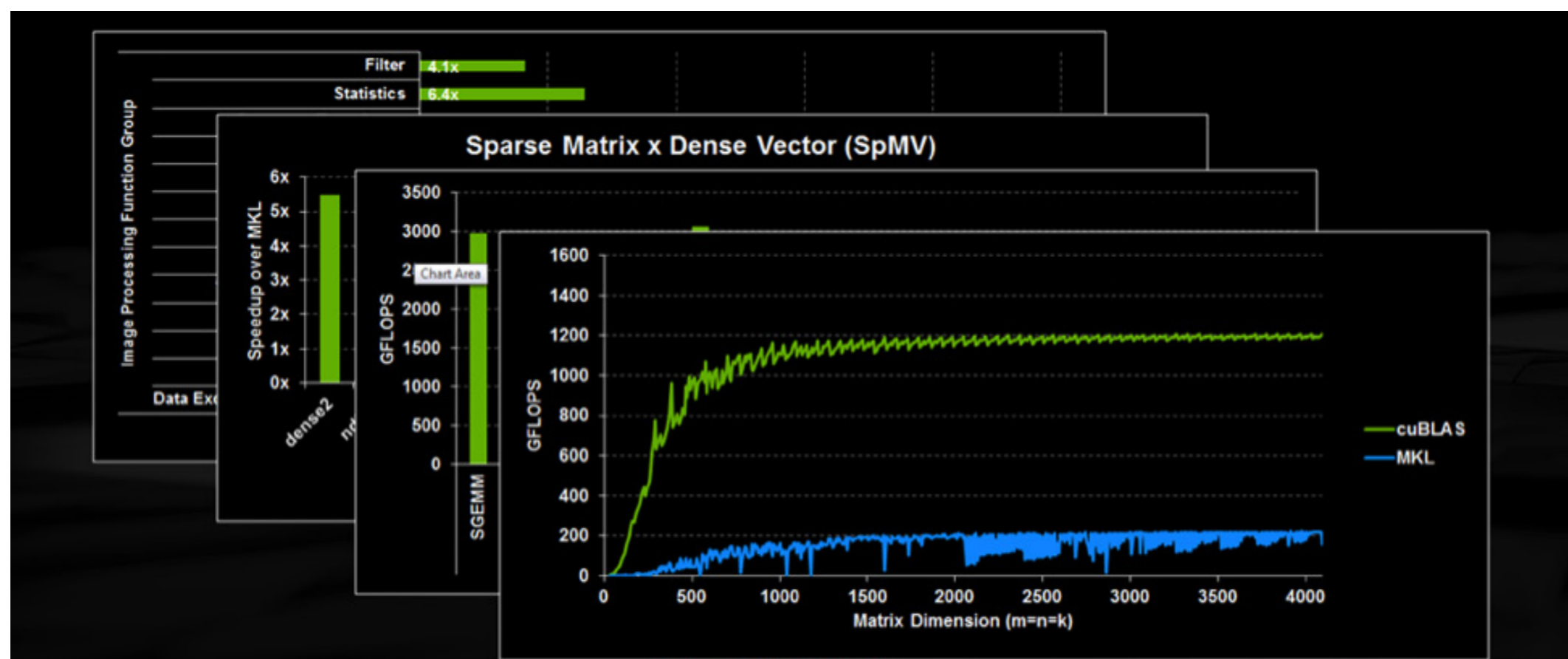


СУПЕРКОМПЬЮТЕРНЫЙ
КОНСОРЦИУМ УНИВЕРСИТЕТОВ РОССИИ

ОТЧЕТ О ПРОИЗВОДИТЕЛЬНОСТИ CUDA 6.5

Доступен отчет о производительности CUDA 6.5 с использованием библиотек:

CUDART | cuFFT | cuBLAS | cuSPARSE | cuRAND | NPP | Thrust | math.h | cuDNN



[Подробнее на английском языке](#)

СТАНЬ ЧАСТЬЮ GTC 2015

У вас есть возможность принять участие в GPU Technology Conference 2015, 17-20 марта 2015 года в Силиконовой Долине США.

Участники из более 50 стран мира приедут на Конференцию по GPU-технологиям, чтобы узнать о потрясающих возможностях параллельных вычислений.

Посетите данное мероприятие вместе с другими учеными, разработчиками, графическими художниками, конструкторами, исследователями, инженерами и IT-менеджерами, чтобы:

- Рассказать о своем научном проекте
- Получить информацию о передовых инновациях
- Узнать о значительных успехах использования GPU
- Поучиться у профессионалов в области параллельных вычислений и визуализации
- Познакомиться с коллегами и экспертами в отрасли

[ПОДАТЬ ЗАЯВКУ НА УЧАСТИЕ](#)

ВЕБИНАРЫ НА АНГЛИЙСКОМ ЯЗЫКЕ

8 октября

**Enabling Next Generation
Intelligent Applications,**
M. Zeiler, Clarifai

ЗАРЕГИСТРИРОВАТЬСЯ

15 октября

Essential CUDA Optimization Techniques,
C. Mason, Acceleware

ЗАРЕГИСТРИРОВАТЬСЯ

НАУЧНЫЕ РАБОТЫ С ИСПОЛЬЗОВАНИЕМ ВЫЧИСЛЕНИЙ НА CUDA

МОДЕЛЬНО-ОРИЕНТИРОВАННЫЙ ПОДХОД К СКВОЗНОМУ ПРОЕКТИРОВАНИЮ АДАПТИВНЫХ РАСПРЕДЕЛЕННЫХ СЕНСОРНЫХ СЕТЕЙ МЕДИЦИНСКОГО НАЗНАЧЕНИЯ.

Володин К.И., Переходов А.И.

В процессе сквозного проектирования адаптивных распределенных сенсорных сетей для медицинского применения, необходимым условием является сопровождение технически сложного проекта, включающего такие аспекты разработки как создание последовательных алгоритмов с последующей их адаптацией для параллельного и распределенного выполнения, верификация полученных алгоритмов путем создания моделей, реализующих запуск алгоритмов на ряде тестовых случаев, моделирование работы адаптивной распределенной сенсорной сети на уровне пакетной передачи, переход к модели, учитывающей специфику медицинского применения цифровых систем (гальваническая изоляция точек обмена информацией, отсутствие или минимизация пропущенных значений), верификации модели, включающей алгоритм экстраполяции по известным значениям, модификация модели для частичной ко-симуляции с физическим оборудованием и последующим переносом на реальное оборудование. Для сопровождения подобного проекта требуется система комплексно реализующая вышеуказанные аспекты.

Методы исследований. В качестве такой системы авторам представляется адекватным использование MathWorks MATLAB&Simulink, а также ряд пакетов расширений (toolboxes) для MATLAB&Simulink, таких как Instrument Control Toolbox, MATLAB Coder, Simulink Coder, Embedded Coder и другие. Алгоритмический аспект может быть смоделирован в MATLAB без применения дополнительных пакетов расширений, позволяя протестировать и визуализировать результаты вычислений. Для дальнейшего тестирования алгоритма в параллельном варианте необходим Parallel Computing Toolbox, позволяющий использовать до 12 ядер одновременно на одной физической машине (версия R2013b). Дальнейшее движение по данному вопросу приводит нас к необходимости использования

решений, позволяющих реализовать кластерные вычисления, демонстрирующие линейный рост скорости вычислений при наращивании количества процессоров, а также необходимости интеграции в процесс моделирования технологии NVIDIA CUDA, предоставляющей ресурс массивного параллелизма в задачах фильтрации, вейвлет преобразования, преобразования Фурье сигналов биологических объектов многомерного характера и т.п.

Модельный аспект характеризуется необходимостью автоматизированного синтеза моделей передачи информации на пакетном уровне, учитывающих специфику медицинского применения. Для создания таких моделей возможно использовать Simulink StateFlow диаграммы, а также SimEvents - расширения блоков Simulink, в целом позволяющие смоделировать сложную систему в терминах системы массового обслуживания (СМО). Дополнительная сложность возникает при необходимости синтеза моделей в автоматизированном режиме при моделировании большой системы (более 1000 элементов). При этом возможно использовать API MATLAB&Simulink для реализации синтезатора модели по входным требованиям, включающим такие ограничения как количество параметров мониторинга и контроля объекта, наличие или отсутствие у узлов адаптивной распределенной сенсорной сети медицинского назначения географической привязки, специализированных тегов и т.п. Поскольку модельный эксперимент обладает определенной долей приближения к реальному объекту, то необходимо протестировать полученные модельные выкладки на реальном оборудовании в режиме ко-симуляции и/или PIL-тестирования. Решение подобной задачи возможно при использовании ряда пакетов расширения MATLAB&Simulink, а также использование т.н. буферной и целевой аппаратной платформы Arduino и сенсорной платформы на базе технологий Nordic Semiconductor соответственно. Метаданные для модельного эксперимента могут быть получены приборами Agilent, Rigol. [ПОДРОБНЕЕ](#)

НАУЧНЫЕ РАБОТЫ С ИСПОЛЬЗОВАНИЕМ ВЫЧИСЛЕНИЙ НА CUDA

ИСПОЛЬЗОВАНИЕ ГПУ ПРОЦЕССОРОВ ДЛЯ КЛАССИФИКАЦИИ ЗАПРОСОВ ПРИ DDoS-АТАКЕ.

А.И. Рейбандт, С.А. Сартин

Новые технологии всё глубже входят во все сферы деятельности человека, системы торговли, документооборота, связи, и даже обеспечения жизнедеятельности всё чаще базируются на кибер платформах. В связи с этим наиболее остро встаёт вопрос о надёжности работы и безопасности систем. Каждый день появляются всё новые и новые типы угроз, взломы и атаки давно перестали носить хаотический характер, хакеры уже не действуют по одиночке, образуя преступную группу с целью обогащения.

Атаки типа DDoS, появились сравнительно недавно, однако быстро стали простым и эффективным инструментом постоянно возрастающей мощностью. Отдельные группы мошенников постоянно изменяют механизмы атак с увеличением зловредного действия, ориентируясь на существующие механизмы защиты. Создаваемый инструментарий в своих алгоритмах использует многоуровневый контроль с шифрованием на каждом участке, тем самым создавая огромные трудности по поимке организатора атаки, что в сочетании с отсутствием в законодательстве Республики Казахстан, а так же стран СНГ, статей наказания, в последнее время привело к бурному росту случаев мошенничества в этом направлении.

Существующие сетевые устройства и традиционные технологии периметрической защиты, в частности, межсетевые экраны и системы выявления вторжений, хотя и служат важными составляющими стратегии безопасности в целом, не обеспечивают полномасштабную защиту от DDoS.

По данным Arbor Networks Inc.[1] мощность DDoS-атак возрастает в среднем в 2 раза за год и на данный момент может достигать мощности 100 Гбит/сек., а выявленные сети зараженных компьютеров – ботнеты превышали сотни тысяч:

Kraken - 400 тысяч компьютеров.

Srizbi - 315 тысяч компьютеров.

Bobax - 185 тысяч компьютеров.

Rustock - 150 тысяч компьютеров.

Storm - 100 тысяч компьютеров.

Psybot - 100 тысяч ADSL-маршрутизаторов, основанных на Linux.

Понятно, что для обработки запросов даже небольшого по размерам ботнета требуются колоссальные мощности сервера, в то время как процесс анализа и адекватности каждого подключения, займёт на порядок больше ресурсов, в убыток аудитории правомерных пользователей, а возможно и ограничит доступ на какое-то время, что в условиях современного рынка просто не приемлемо. Покупка сервиса защиты от DDoS, наращивание мощностей сервера с усложнением конфигурации являются хоть и эффективными, но всё же пока ещё дорогими способами противодействия злоумышленникам.

Предлагаемый подход является более простым и дешевым решением, основанным на возможности независимой обработки каждого подключения и как следствие достижения большой степени параллелизма. Алгоритм распознавания DDoS-атаки среди пользовательских запросов имеет ряд усовершенствований, отличающих его от уже существующих подходов. [ПОДРОБНЕЕ](#)

НАУЧНЫЕ РАБОТЫ С ИСПОЛЬЗОВАНИЕМ ВЫЧИСЛЕНИЙ НА CUDA

ПАРАЛЛЕЛЬНАЯ ВЕРСИЯ ЖАДНОГО АЛГОРИТМА КЛАСТЕРИЗАЦИИ

М.А. Баранов

Одной из наиболее важных в информационном поиске задач является кластеризация – разбиение исходного множества объектов на группы, состоящие из схожих объектов. Кластеризация нашла широкое применение в различных областях знаний: в биологии, социологии, информатике, астрономии, медицине, археологии, маркетинговых исследованиях.

Кластерный анализ является примером задачи обучения без учителя, его основная цель – разбить множество объектов на группы таким образом, чтобы объекты внутри одной группы были максимально похожими друг на друга, но в то же время максимально отличались от объектов другой группы.

В настоящее время разработано множество алгоритмов кластеризации, использующих различные подходы к решению задачи кластерного анализа. Из всего многообразия используемых

ОБЗОР МЕТОДОВ УПРОЩЕНИЯ ПОЛИГОНАЛЬНЫХ МОДЕЛЕЙ НА ГРАФИЧЕСКОМ ПРОЦЕССОРЕ

А. В. Хоперсков, С. А. Хоперсков, С. С. Храпов, А. В. Мишустин, А. В. Моисеев

Численное моделирование на основе суперкомпьютерных технологий является мощным инструментом изучения сложных динамических процессов в галактиках, предоставляя возможность учитывать богатую физику в многокомпонентных галактических системах.

Приведены результаты численного моделирования галактик с целью определения свойств темного вещества внутри 2-3 оптических радиусов галактики. Расчеты основаны на бесстолкновительных моделях N-тел, трехмерных газодинамических кодах и самосогласованных звездно-газовых моделях.

при кластеризации подходов стоит выделить так называемые “жадные” алгоритмы, суть работы которых сводится к тому, что на каждом шаге они делают локально оптимальный выбор в расчете на то, что это приведет к оптимальному решению всей задачи. Стоит отметить, что жадные алгоритмы часто используются при решении задач кластеризации.

Целью данной работы является сравнительный анализ параллельной и последовательной версий одного из жадных алгоритмов кластеризации при обработке коллекций документов разумного размера (несколько тысяч документов). Для решения задачи распараллеливания алгоритма была выбрана технология CUDA, разработанная компанией NVIDIA.

По результатам исследования был сделан вывод о том, что при кластеризации коллекций, состоящих из более чем 5000 документов, параллельная версия алгоритма кластеризации работает быстрее последовательной версии, причем различие во времени выполнения становится тем больше, чем больше в коллекции документов. [ПОДРОБНЕЕ](#)

Обсуждаются особенности реализации численных моделей для параллельных вычислений на суперкомпьютерах с массивно-параллельной архитектурой и гибридных вычислительных системах с графическими ускорителями.

В работе дан краткий обзор результатов моделирования галактик с целью получения ограничений на параметры распределений темного вещества внутри конкретных дисковых галактик и в ближайшей их окрестности. Основное внимание уделено на особенности реализации численных моделей для параллельных вычислений на суперкомпьютерах с массивно-параллельной архитектурой и гибридных вычислительных системах с графическими ускорителями с использованием технологии CUDA. [ПОДРОБНЕЕ](#)

ПОЛЕЗНЫЕ РЕСУРСЫ ПО CUDA

Форум Разработчиков NVIDIA

Присоединяйтесь к Форуму CUDA-разработчиков, делитесь своим опытом и узнавайте много нового. <http://devtalk.nvidia.com/>

Документация по CUDA

Со списком документации по CUDA можно ознакомиться [здесь](#).

Обучение онлайн

[Udacity](#) | [Coursera](#) | [Курс на русском языке](#)

Библиотеки с поддержкой GPU ускорения

Список библиотек с поддержкой GPU ускорения от NVIDIA и партнеров.

<https://developer.nvidia.com/gpu-accelerated-libraries>

GPU Тест-Драйв

Хотите бесплатно протестировать Tesla K40? Зарегистрируйтесь [здесь](#).

Ускоряйте научные приложения с OpenACC

Протестируйте компилятор PGI OpenACC бесплатно в течение месяца. [Подробнее](#).

Приложения, ускоряемые на GPU

Ознакомьтесь со списком из более 270 приложений [можно на сайте](#).

Книги, посвященные CUDA и вычислениям на GPU

Со списком книг, посвященных CUDA и вычислениям на GPU, можно ознакомиться [здесь](#).

Скачайте

CUDA <http://developer.nvidia.com/cuda-downloads>

Nsight <http://www.nvidia.com/object/nsight.html>

ВАКАНСИИ CUDA

Ведущий программист C++ (3D алгоритмы). Адалиск. Москва.

Стабильной компании специализирующейся на решении задач связанных с трехмерным моделированием и обработкой 3D данных, давно и успешно работающей с крупным стратегическим партнером в США, требуется Ведущий программист C++ для разработки 3D CAD системы.

Обязанности:

- Разработка CAD системы.
- Создание новых и усовершенствование существующих алгоритмов для работы с 3D моделями .
- Написание кода инфраструктурных классов.
- Поддержание и доработка существующего кода.

Требования:

- Опыт разработки на C++ от 5 лет, знание современных подходов к построению систем на C++.
- Умение читать и разбираться в чужом C++ коде
- Хорошая математическая подготовка (желательно физическое или математическое образование), знание предметной области и алгоритмов аналитической геометрии и линейной алгебры, знание вычислительной математики.
- Знание английского языка на уровне необходимом для ведения переписки.
- Желательно знание технологии NVidia CUDA.
- Эта работа для тех, кто хорошо знает C++ и не чужд математики.
- Кто любит писать быстрый, хорошо оптимизированный код.
- Кому нравится изящество C++. Если вы действительно научились в университете чему-то полезному, а сейчас понимаете, что в вашей работе эти знания не требуются, если вам не хочется заниматься базами данных и ваянием отчетов, если вы не хотите чтобы ваши знания исчезали за ежедневной рутинной, то наша позиция для вас.

Условия:

- Гибкий график. Оформление по ТК.
- Бесплатные обеды в офисе.
- Комфортные условия работы: современная техника, офис с кондиционерами, столом для пинг-понга и т.д.
- Дружный доброжелательный коллектив, в котором приятно работать.
- Возможны командировки в США.
- Чего точно не будет: вы не будете заниматься базами данных, написанием формочек

[ПОДРОБНЕЕ](#)

ВАКАНСИИ CUDA

PhysX software engineer. NVIDIA. Москва.

NVIDIA is searching for world-class software engineers for an exciting role in our Tegra game developer technology team. Work with the most advanced mobile computing technology as the world shifts to mobile and green computing. Interact closely with the architecture and software teams at NVIDIA, as well as external game developers to ensure the best possible performance and results. Work to bring world-class games to the Tegra platform.

- Strong knowledge of C/C++ and programming techniques
- Strong mathematical fundamentals
- Excellent communication skills required
- Travel for on-site visits with developers and to conferences may be required
- Ideal candidates will have experience with game development, DirectX/OpenGL, Windows and Linux
- Minimum 3 years of industry or academic experience (or equivalent) in a related field
- B.S. or higher degree in Computer Science/Engineering or mathematical field
- Experience with Java and Android a plus
- Experience with CUDA, DirectX, or OpenGL ES is a plus
- A strong team player that is self motivated is a requirement

[ПОДРОБНЕЕ](#)

КОНТАКТЫ

Если вы хотите, чтобы ваша статья появилась в следующем выпуске CUDA Альманах пишите нам на:

Лидия Андреева
landreeva@nvidia.com

По вопросам приобретения NVIDIA GPU и по прочим техническим вопросам пишите нам на:

Антон Джораев
adzhoraev@nvidia.com