

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Институт вычислительной математики и информационных технологий



подписано электронно-цифровой подписью

Программа дисциплины
Высокопроизводительные вычисления БЗ.ДВ.6

Направление подготовки: 010400.62 - Прикладная математика и информатика

Профиль подготовки: Численные методы

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Автор(ы):

Кадыров Р.Ф.

Рецензент(ы):

Задворнов О.А.

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Задворнов О. А.

Протокол заседания кафедры No ____ от " ____ " _____ 201__ г

Учебно-методическая комиссия Института вычислительной математики и информационных технологий:

Протокол заседания УМК No ____ от " ____ " _____ 201__ г

Регистрационный No 9114014

Казань
2014

Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) ассистент, к.н. Кадыров Р.Ф. кафедра вычислительной математики отделение прикладной математики и информатики, Rafael.Kadyrov@kpfu.ru

1. Цели освоения дисциплины

Целью дисциплины является изучение основных архитектур многопроцессорных систем и принципов разработки прикладного программного обеспечения для них.

2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б3.ДВ.6 Профессиональный" основной образовательной программы 010400.62 Прикладная математика и информатика и относится к дисциплинам по выбору. Осваивается на 4 курсе, 7 семестр.

Данная дисциплина относится к общепрофессиональным дисциплинам.

Читается на 4 курсе в 7 семестре для студентов обучающихся по направлению "Прикладная математика и информатика".

3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-8 (профессиональные компетенции)	способность разрабатывать наукоемкое программное обеспечение работы конкретного предприятия
ПК-9 (профессиональные компетенции)	способность и готовностью проводить научные эксперименты, оценивать результаты исследований

В результате освоения дисциплины студент:

4. должен продемонстрировать способность и готовность:

В результате изучения курса "Программирование в среде много процессорных комплексов" студент приобретает следующие знания, навыки и умения:

1. знание архитектуры современных математических и графических сопроцессоров
2. способность самостоятельно понять и изучить архитектуру вновь появляющихся ускорителей
3. знание принципов разработки ПО для современных GPU
4. ориентироваться в стеке технологий CUDA для GPU nVidia
5. знать состав библиотеки CUDA для неграфических вычислений, уметь применять эти библиотеки при разработке ПО для GPU nVidia

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 3 зачетных(ые) единиц(ы) 108 часа(ов).

Форма промежуточного контроля дисциплины зачет в 7 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);
 55-70 баллов - "удовлетворительно" (удов.);
 54 балла и менее - "неудовлетворительно" (неуд.).

4.1 Структура и содержание аудиторной работы по дисциплине/ модулю Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Графический конвейер. Архитектура GPU-устройства.	6		0	0	6	устный опрос
2.	Тема 2. Иерархия памяти GPU. Константная, глобальная, текстурная, разделяемая память.	6		0	0	6	устный опрос
3.	Тема 3. Общие принципы построения программ для GPU. Модель программирования в общей памяти. SIMD (SIMT) модель программы.	6		0	0	6	устный опрос
4.	Тема 4. Программная модель CUDA.	6		0	0	6	коллоквиум
5.	Тема 5. Некоторые алгоритмы обработки массивов. Параллельная редукция. Префиксная сумма.	6		0	0	6	контрольная работа домашнее задание
6.	Тема 6. Некоторые численные алгоритмы.	6		0	0	8	контрольная работа домашнее задание
7.	Тема 7. Прикладные математические библиотеки: CUBLAS, CUSPARSE, CUFFT, CURAND.	6		0	0	8	коллоквиум домашнее задание

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
8.	Тема 8. Высокоуровневые технологии разработки. Введение в Thrust. Реализация вычисления числа "пи" составными квадратурными формулами при помощи Thrust. Сравнение производительности.	6		0	0	8	контрольная работа домашнее задание
	Тема . Итоговая форма контроля	7		0	0	0	зачет
	Итого			0	0	54	

4.2 Содержание дисциплины

Тема 1. Графический конвейер. Архитектура GPU-устройства.

лабораторная работа (6 часа(ов)):

Эволюция графических ускорителей. Появление и эволюция графического конвейера. Текстуры, шейдеры. Шейдерные процессоры. Общие черты внутреннего устройства графических ускорителей разных производителей.

Тема 2. Иерархия памяти GPU. Константная, глобальная, текстурная, разделяемая память.

лабораторная работа (6 часа(ов)):

Виды памяти в GPU устройствах. Константная, глобальная, текстурная, разделяемая память. Расположение на графическом чипе, кеширование, ограничения доступа. Обмен данными между GPU и CPU. Общее виртуальное адресное пространство.

Тема 3. Общие принципы построения программ для GPU. Модель программирования в общей памяти. SIMD (SIMT) модель программы.

лабораторная работа (6 часа(ов)):

Классификация Флинна. Место GPU в классификации Флинна. Отличия модели SIMT от классической SIMD-архитектуры.

Тема 4. Программная модель CUDA.

лабораторная работа (6 часа(ов)):

Программная модель CUDA. Взаимодействие CPU->GPU->CPU. Взаимодействие CUDA и C/C++. Расширение языка C/C++. Встроенные типы данных, дополнительные языковые конструкции. CUDA-библиотека времени исполнения. Атомарные операции.

Тема 5. Некоторые алгоритмы обработки массивов. Параллельная редукция.

Префиксная сумма.

лабораторная работа (6 часа(ов)):

Установка и настройка программного обеспечения CUDA под ОС семейств Windows и Linux. Расширение языка C/C++. Встроенные типы данных, дополнительные языковые конструкции. CUDA-библиотека времени исполнения. Атомарные операции. Компиляция CUDA-программ.

Тема 6. Некоторые численные алгоритмы.

лабораторная работа (8 часа(ов)):

Программная реализация алгоритма параллельного суммирование элементов одномерного массива на GPU. Сравнение производительности CPU и GPU-реализаций.

Тема 7. Прикладные математические библиотеки: CUBLAS, CUSPARSE, CUFFT, CURAND.

лабораторная работа (8 часа(ов)):

Программная реализация на GPU алгоритмов: - транспонирования матрицы - вычисления числа "пи" при помощи составных квадратурных формул - вычисления числа "пи" методом Монте-Карло.

Тема 8. Высокоуровневые технологии разработки. Введение в Thrust. Реализация вычисления числа "пи" составными квадратурными формулами при помощи Thrust. Сравнение производительности.

лабораторная работа (8 часа(ов)):

Введение в шаблоны C++. Функтор, итератор. Их реализация на C++. ZIP-итератор. Общая идеология Thrust. Примеры использования. Взаимодействие Thrust и CUDA Plain C.

4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
1.	Тема 1. Графический конвейер. Архитектура GPU-устройства.	6		подготовка к устному опросу	6	устный опрос
2.	Тема 2. Иерархия памяти GPU. Константная, глобальная, текстурная, разделяемая память.	6		подготовка к устному опросу	6	устный опрос
3.	Тема 3. Общие принципы построения программ для GPU. Модель программирования в общей памяти. SIMD (SIMT) модель программы.	6		подготовка к устному опросу	6	устный опрос
4.	Тема 4. Программная модель CUDA.	6		подготовка к коллоквиуму	3	коллоквиум
				подготовка к устному опросу	3	устный опрос
5.	Тема 5. Некоторые алгоритмы обработки массивов. Параллельная редукция. Префиксная сумма.	6		подготовка домашнего задания	4	домашнее задание
				подготовка к контрольной работе	2	контрольная работа

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
6.	Тема 6. Некоторые численные алгоритмы.	6		подготовка домашнего задания	4	домашнее задание
				подготовка к контрольной работе	4	контрольная работа
7.	Тема 7. Прикладные математические библиотеки: CUBLAS, CUSPARSE, CUFFT, CURAND.	6		подготовка домашнего задания	2	домашнее задание
				подготовка к коллоквиуму	6	коллоквиум
8.	Тема 8. Высокоуровневые технологии разработки. Введение в Thrust. Реализация вычисления числа "пи" составными квадратурными формулами при помощи Thrust. Сравнение производительности.	6		подготовка домашнего задания	2	домашнее задание
				подготовка к контрольной работе	3	контрольная работа
				подготовка к контрольной работе	3	контрольная точка
Итого					54	

5. Образовательные технологии, включая интерактивные формы обучения

Обучение происходит в форме лекционных, лабораторных занятий, а также самостоятельной работы студентов.

Изучение курса подразумевает получение практических навыков для более глубокого понимания разделов дисциплины на основе решения задач и упражнений, иллюстрирующих доказываемые теоретические положения, а также развитие абстрактного мышления и способности самостоятельно доказывать частные утверждения.

Самостоятельная работа предполагает выполнение домашних работ. Практические задания, выполненные в аудитории, предназначены для указания общих методов решения задач определенного типа. Закрепить навыки можно лишь в результате самостоятельной работы.

Кроме того, самостоятельная работа включает подготовку к экзамену. При подготовке к сдаче экзамена весь объем работы рекомендуется распределять равномерно по дням, отведенным для подготовки к экзамену, контролировать каждый день выполнения работы. Лучше, если можно перевыполнить план. Тогда всегда будет резерв времени.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

Тема 1. Графический конвейер. Архитектура GPU-устройства.

устный опрос , примерные вопросы:

Появление и эволюция графического конвейера. Текстуры, шейдеры. Шейдерные процессоры. Общие черты внутреннего устройства графических ускорителей разных производителей.

Тема 2. Иерархия памяти GPU. Константная, глобальная, текстурная, разделяемая память.

устный опрос , примерные вопросы:

Константная, глобальная, текстурная, разделяемая память. Расположение на графическом чипе, кеширование, ограничения доступа. Обмен данными между GPU и CPU. Общее виртуальное адресное пространство.

Тема 3. Общие принципы построения программ для GPU. Модель программирования в общей памяти. SIMD (SIMT) модель программы.

устный опрос , примерные вопросы:

Классификация Флинна. Место GPU в классификации Флинна. Отличия модели SIMT от классической SIMD-архитектуры.

Тема 4. Программная модель CUDA.

коллоквиум , примерные вопросы:

Программная модель CUDA. Взаимодействие CPU->GPU->CPU. Взаимодействие CUDA и C/C++.

устный опрос , примерные вопросы:

Расширение языка C/C++. Встроенные типы данных, дополнительные языковые конструкции. CUDA-библиотека времени исполнения. Атомарные операции. Компиляция CUDA-программ.

Тема 5. Некоторые алгоритмы обработки массивов. Параллельная редукция. Префиксная сумма.

домашнее задание , примерные вопросы:

Программная реализация алгоритма префиксного суммирование элементов одномерного массива на GPU.

контрольная работа , примерные вопросы:

- архитектура GPU nVidia - графический конвейер

Тема 6. Некоторые численные алгоритмы.

домашнее задание , примерные вопросы:

Вычисление числа "пи" методами квадратурных формул и Монте Карло.

контрольная работа , примерные вопросы:

- виды памяти GPU - программная модель CUDA

Тема 7. Прикладные математические библиотеки: CUBLAS, CUSPARSE, CUFFT, CURAND.

домашнее задание , примерные вопросы:

Сравнение точности метода Монте-Карло при использовании различных генераторов псевдослучайных последовательностей из библиотеки CURAND

коллоквиум , примерные вопросы:

Способы представления разреженных матриц в программе. Реализация основных матричных и матрично-векторных операций линейной алгебры при том или ином подходе к программному представлению матрицы.

Тема 8. Высокоуровневые технологии разработки. Введение в Thrust. Реализация вычисления числа "пи" составными квадратурными формулами при помощи Thrust.

Сравнение производительности.

домашнее задание , примерные вопросы:

Реализация префиксного суммирования при помощи библиотеке Thrust. Сравнение вычислительной эффективности с низкоуровневой реализацией алгоритма.

контрольная работа , примерные вопросы:

- структура библиотеки CUBLAS - структура библиотеки CUSPARSE

контрольная точка , примерные вопросы:

- структура библиотеки CURAND - идеология формулировки алгоритмов при помощи библиотеки Thrust - функтор, итератор, zip-итератор

Тема . Итоговая форма контроля

Примерные вопросы к зачету:

Список вопросов:

1. Эволюция GPU. Современный графический конвейер.
2. Архитектура графического ускорителя nVidia. Шейдер. Шейдерный процессор. Управление потоком инструкций и данных.
3. Виды памяти GPU. Константная, глобальная, текстурная, разделяемая память. Расположение на графическом чипе, кеширование, ограничения доступа.
4. Программная модель CUDA. Расширение языка C/C++. Встроенные типы данных, дополнительные языковые конструкции.
5. Программная модель CUDA. Расширение языка C/C++. CUDA-библиотека времени исполнения. Атомарные операции.
6. Предназначение и структура библиотек CUBLAS, CUSPARSE, CURAND.
7. Библиотека Thrust. Идеология. Детали реализации вычислительных алгоритмов с использованием Thrust. Функтор, итератор, zip-итератор.
8. Опишите структуру графического процессора G80.
9. Каковы основные вычислительные возможности нитевых ядер?
10. Опишите структуру потокового мультипроцессора.
11. Как должны быть организованы вычисления в процессоре G80?
12. В чем состоят принципы организации и функционирования массива взаимодействующих нитей?
13. В соответствии с техникой массива взаимодействующих нитей распишите один из известных вам алгоритмов численного интегрирования.

7.1. Основная литература:

1. Линев, Алексей Владимирович. Технологии параллельного программирования для процессоров новых архитектур: учебник для студентов высших учебных заведений/ А. В. Линев, Д. К. Боголепов, С. И. Бастраков; под ред. В. П. Гергеля; Нижегород. гос. ун-т им. Н. И. Лобачевского. Москва: Изд-во Московского университета, 2010. 148 с.
2. Инструменты параллельного программирования в системах с общей памятью: учебник для студентов высших учебных заведений/ К.В. Корняков, В.Д. Кустикова, И.Б. Мееров [и др.]; под ред. проф. В.П. Гергеля; Нижегород. гос. ун-т им. Н.И. Лобачевского, Координац. совет Системы науч.-образоват. центров суперкомпьютер. технологий. 2-е изд., испр. и доп.. Москва: Изд-во Московского университета, 2010. 262 с.
3. Гергель, Виктор Павлович. Высокопроизводительные вычисления для многопроцессорных многоядерных систем : учебник для студентов высших учебных заведений, обучающихся по направлениям ВПО 010400 "Прикладная математика и информатика" и 010300 "Фундаментальная информатика и информационные технологии" / В. П. Гергель ; Б-ка Нижегород. гос. ун-та им. Н. И. Лобачевского .? Москва : Изд-во Московского университета [и др.], 2010 .? 539, [4] с.
3. Сырецкий, Г. А. Информатика. Фундаментальный курс. Том II. Информационные технологии и системы / Г. А. Сырецкий. ? СПб.: БХВ-Петербург, 2007. ? 846 с.
<http://znanium.com/bookread.php?book=350042>
4. Максимов Н. В. Технические средства информатизации: Учебник / Н.В. Максимов, Т.Л. Партыка, И.И. Попов. - 4-е изд., перераб. и доп. - М.: Форум: НИЦ ИНФРА-М, 2013. - 608 с.:
<http://znanium.com/bookread.php?book=410390>
5. Богачёв К.Ю. Основы параллельного программирования: учебное пособие. - 2-е (эл.). - М.: Бином. Лаборатория знаний, 2013. - 342 с.
http://e.lanbook.com/books/element.php?pl1_id=42626

7.2. Дополнительная литература:

1. Ефимов, С. С. Параллельное программирование: учебное пособие / С. С. Ефимов; Федер. агентство по образованию, ОмГУ, Фак. компьютер. наук. Омск: [УниПак], 2009. 397 с

2. Масловская Л. В.. Параллельные алгоритмы: учебное пособие для студентов и аспирантов университетов и вузов, изучающих вычислительную математику и ее приложения, а также для специалистов по численному анализу / Л. В. Масловская, О. М. Масловская. - Одесса: Фенікс, 2009. - 109 с.:

3. Столов Е. Л. Введение в цифровую обработку изображений и параллельные вычисления: [учеб. пособие] / Е. Л. Столов; Казан. гос. ун-т. - Казань: [КГУ], 2006. - 67, [1] с.: ил.; 20. - Библиогр.: с. 68 (3 назв.).

7.3. Интернет-ресурсы:

Интернет-портал образовательных ресурсов по ИТ - <http://algotlist.manual.ru/>

Интернет-портал образовательных ресурсов по ИТ - <http://www.intuit.ru>

Информационно-аналитический центр по параллельным вычислениям - <http://parallel.ru>

Основы работы с технологией CUDA - <http://www.znaniium.com>

Параллельные алгоритмы: учебное пособие для студентов и аспирантов университетов и вузов, изучающих вычислительную математику и ее приложения, а также для специалистов по численному анализу - http://z3950.ksu.ru/bcover/0000730145_con.pdf

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Высокопроизводительные вычисления" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

Лекционные занятия проводятся в аудитории, оснащенной доской и мелом (маркером). Желательно наличие мультимедийного проектора.

Практические занятия проводятся в компьютерных классах, оснащенных доской и мелом (маркером). ЭВМ должны быть оснащены видеоадаптерами nVidia с поддержкой технологии CUDA.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 010400.62 "Прикладная математика и информатика" и профилю подготовки Численные методы .

Автор(ы):

Кадыров Р.Ф. _____

"__" _____ 201__ г.

Рецензент(ы):

Задворнов О.А. _____

"__" _____ 201__ г.