

МИНОБРНАУКИ РОССИИ
Ярославский государственный университет им. П.Г. Демидова

Кафедра математического моделирования

УТВЕРЖДАЮ

Декан математического факультета

Нестеров П.Н.

21 мая 2024 г.

Рабочая программа дисциплины
Технологии многопоточных вычислений

Направление подготовки (специальности)
01.04.02 Прикладная математика и информатика

Направленность (профиль)
«Математическое моделирование и численные методы»

Форма обучения очная

Программа рассмотрена
на заседании кафедры
от 12.04.2024, протокол № 8

Программа одобрена НМК
математического факультета
протокол № 9 от 03.05.2024

1. Цели освоения дисциплины

Целями освоения дисциплины «Технологии многопоточных вычислений» являются основы параллельных вычислений на графических адаптерах NVIDIA. Данный курс вырабатывает у студентов навыки как параллельного программирования на GPGPU так и работы с распределенными системами.

2. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии многопоточных вычислений» относится к обязательной части образовательной программы. Данная дисциплина предоставляет базовые навыки, необходимые специалистам работающих в сфере машинного обучения и больших данных.

Для освоения данной дисциплиной студенты должны владеть математическим аппаратом векторного и тензорного анализа, линейной алгебры, дифференциального исчисления, уметь решать основные задачи численных методов, а также иметь навыки программирования на языке C++.

Полученные в курсе «Технологии многопоточных вычислений» знания необходимы специалистам, разрабатывающим высокопроизводительные приложения и распределённые системы.

3. Планируемые результаты обучения по дисциплине, соотнесённые с планируемыми результатами освоения образовательной программы

Процесс изучения дисциплины направлен на формирование следующих компетенций в соответствии с ФГОС ВО, ООП ВО и приобретения следующих знаний, умений, навыков и (или) опыта деятельности:

Формируемая компетенция (код и формулировка)	Индикатор достижения компетенции (код и формулировка)	Перечень планируемых результатов обучения
Общепрофессиональные компетенции		
ОПК-4 Способен комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	ИД-ОПК-4.1 Обладает обширными знаниями в области информационных технологий, применяемых в профессиональной деятельности	Знать: – о предпосылках рождения параллельных вычислений и GPGPU. – устройство и архитектуру GPU; – механизмы взаимодействия CPU и GPU. – Особенности CUDA Уметь: – выбирать правильные средства оптимизации, исходя из контекста задачи; – масштабировать решения для различных программно-аппаратных архитектур; – работать с распределёнными системами Владеть навыками: – написания параллельных программ; – разработки параллельных алгоритмов – параллельных вычислений

4. Объём, структура и содержание дисциплины

Общая трудоёмкость дисциплины составляет **2** зачётных единицы, **72** акад. часа.

№ п/п	Темы (разделы) дисциплины, их содержание	Семестр	Виды учебных занятий, включая самостоятельную работу студентов, и их трудоемкость (в академических часах)						Формы текущего контроля успеваемости Форма промежуточной аттестации (по семестрам)
			Контактная работа						
			лекции	практические	лабораторные	консультации	аттестационные испытания	самостоятельная работа	
1	Вводная лекция	1	2					10	Задания для самостоятельной работы
2	Архитектура	1	2					10	Задания для самостоятельной работы
3	Основы CUDA	1	2		3	1		10	Задания для самостоятельной работы, Лабораторная работа №1
4	Работа с памятью	1	1		2			10	Задания для самостоятельной работы
5	Средства оптимизации	1	1		3	1		10	Задания для самостоятельной работы, Лабораторная работа №2
							0,3	3,7	Зачёт
	ИТОГО		8		8	2	0,3	53,7	

Содержание разделов дисциплины:

1. Вводная лекция.

- 1.1. Введение. Обзор технологий.
- 1.2. Гибридная модель вычислений.
- 1.3. Типы вычислительных архитектур.

2. Архитектура.

- 2.1. Архитектура GPU.
- 2.2. CUDA. Первый взгляд.
- 2.3. Типы вычислительных архитектур.

3. Основы CUDA.

- 3.1. Программная модель CUDA.
- 3.2. Гибридная модель программного кода.
- 3.3. Thread, block, grid blocks.

3.4. Kernel - функция.

4. Работа с памятью.

4.1. Типы памяти.

4.2. Единая система памяти (Unified Memory).

4.3. Pinned memory.

4.4. CUDA streams.

5. Средства оптимизации.

5.1. Объединение запросов (Memory coalescing).

5.2. Разделяемая память (Shared memory).

5.3. Оптимизация работы с разделяемой памятью.

5.4. Банк-конфликты.

5. Образовательные технологии, в том числе технологии электронного обучения и дистанционные образовательные технологии, используемые при осуществлении образовательного процесса по дисциплине

В процессе обучения используются следующие образовательные технологии:

Вводная лекция – дает первое целостное представление о дисциплине и ориентирует студента в системе изучения данной дисциплины. Студенты знакомятся с назначением и задачами курса, его ролью и местом в системе учебных дисциплин и в системе подготовки в целом. Дается краткий обзор курса, история развития науки и практики, достижения в этой сфере, имена известных ученых, излагаются перспективные направления исследований. На этой лекции высказываются методические и организационные особенности работы в рамках данной дисциплины, а также дается анализ рекомендуемой учебно-методической литературы.

Академическая лекция с элементами лекции-беседы – последовательное изложение материала, осуществляемое преимущественно в виде монолога преподавателя. Элементы лекции-беседы обеспечивают контакт преподавателя с аудиторией, что позволяет привлекать внимание студентов к наиболее важным темам дисциплины, активно вовлекать их в учебный процесс, контролировать темп изложения учебного материала в зависимости от уровня его восприятия.

Консультации – вид учебных занятий, являющийся одной из форм контроля самостоятельной работы студентов. На консультациях по просьбе студентов рассматриваются наиболее сложные моменты при освоении материала дисциплины, преподаватель отвечает на вопросы студентов, которые возникают у них в процессе самостоятельной работы.

Лабораторная работа – организация учебной работы с реальными материальными и информационными объектами, экспериментальная работа с аналоговыми моделями реальных объектов.

6. Перечень лицензионного и (или) свободно распространяемого программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине

В процессе осуществления образовательного процесса по дисциплине используются:

для формирования материалов для текущего контроля успеваемости и проведения промежуточной аттестации, для формирования методических материалов по дисциплине:

- программы Microsoft Office;
- издательская система LaTeX;

- Adobe Acrobat Reader;
- MikTeX (свободно распространяемое ПО).

7. Перечень современных профессиональных баз данных и информационных справочных систем, используемых при осуществлении образовательного процесса по дисциплине (при необходимости)

В процессе осуществления образовательного процесса по дисциплине используются:

- Автоматизированная библиотечно-информационная система «БУКИ-NEXT»
http://www.lib.uniyar.ac.ru/opac/bk_cat_find.php
- Электронная библиотечная система «Лань» <https://e.lanbook.com>
- Электронная библиотечная система «Юрайт» <https://urait.ru>
- Электронная библиотечная система «Консультант студента»
<https://www.studentlibrary.ru>

8. Перечень основной и дополнительной учебной литературы, ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости), рекомендуемых для освоения дисциплины

а) основная литература

1. Биллиг В. А. Параллельные вычисления и многопоточное программирование - Москва: НОУ "ИНТУИТ", 2016. https://www.studentlibrary.ru/ru/book/intuit_242.html
2. Боресков А. В., Харламов А. А. Основы работы с технологией CUDA - Москва: ДМК Пресс, 2010. <https://www.studentlibrary.ru/ru/book/ISBN9785940745785.html>

б) дополнительная литература

1. Чернышев, С. А. Основы программирования на Python : учебное пособие для вузов / С. А. Чернышев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 349 с. — (Высшее образование). — ISBN 978-5-534-17139-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/544190>

в) ресурсы сети «Интернет»

1. <https://cuda.grandrust.com>
2. <https://www.udacity.com/course/intro-to-parallel-programming--cs344>
3. <http://getcourse.ispras.ru/course/view.php?id=3>

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине включает в свой состав специальные помещения:

- учебные аудитории для проведения занятий лекционного типа;
- учебные аудитории для проведения лабораторных занятий;
- учебные аудитории для проведения групповых и индивидуальных консультаций;
- учебные аудитории для проведения текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы;
- помещения для хранения и профилактического обслуживания технических средств обучения.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде ЯрГУ.

Автор(ы) :

Старший преподаватель
кафедры математического моделирования

Д. О. Логинов

**Приложение №1 к рабочей программе дисциплины
«Технологии многопоточных вычислений»**

**Фонд оценочных средств
для проведения текущего контроля успеваемости
и промежуточной аттестации студентов
по дисциплине**

**1. Типовые контрольные задания и иные материалы,
используемые в процессе текущего контроля успеваемости**

Задания для самостоятельной работы

Задания по теме № 1 «Введение»:

Раздел 1.1: Знакомство с аппаратными средствами. Написать программу, выводящую информацию о текущем GPU карте.

Задания по теме № 2 «Архитектура»:

Раздел 2.2: Написать простейшую программу, работающую на GPU.

Задания по теме № 3 «Основы CUDA»:

Раздел 3.4: Запрограммировать алгоритм параллельного вычисления заданной функции.

Задания по теме № 4 «Работа с памятью»:

Раздел 4.3: Запрограммировать алгоритм параллельного вычисления заданной функции используя pinned память

Раздел 4.4: Запрограммировать алгоритм параллельного вычисления заданной функции используя CUDA потоки

Задания по теме № 5 «Средства оптимизации»:

Раздел 5.3: Запрограммировать алгоритм параллельного вычисления заданной функции используя shared memory

Раздел 5.4: Запрограммировать алгоритм параллельного вычисления заданной функции разрешив банк-конфликты

Лабораторная работа № 1

Реализовать параллельный алгоритм вычисления основных операций векторного пространства используя технологию CUDA (сложение, умножение, умножение на скаляр).

Лабораторная работа № 2

Реализовать параллельный алгоритм, а так же произвести сравнительный анализ, средств оптимизации вычисления произведения матриц используя технологию CUDA.

Приложение №2 к рабочей программе дисциплины «Технологии многопоточных вычислений»

Методические указания для студентов по освоению дисциплины

Основной формой изложения учебного материала по дисциплине являются лекции и практические занятия. Как правило, на первой половине занятия проводится лекция, в которой излагается основной материал, а затем учащимся выдается практическое задание для освоения и закрепления лекционного материала путем применения его к конкретным задачам и отработка навыков работы с технологией CUDA.

Для успешного освоения дисциплины очень важно решение достаточно большого количества задач, как в аудитории, так и самостоятельно в качестве домашних заданий. Примеры решения задач разбираются на лекциях и практических занятиях, при необходимости по наиболее трудным темам проводятся дополнительные консультации. Основная цель решения задач – помочь усвоить основы параллельных вычислений на графических адаптерах. Для решения всех задач необходимо знать и понимать лекционный материал. Поэтому в процессе изучения дисциплины рекомендуется регулярное повторение пройденного лекционного материала. Предоставленный материал необходимо дома еще раз прорабатывать и при необходимости дополнять информацией, полученной на консультациях, практических занятиях или из учебной литературы.

Большое внимание должно быть уделено выполнению домашней работы. В качестве заданий для самостоятельной работы дома студентам предлагаются задачи, аналогичные разобранным на лекциях и практических занятиях или немного более сложные, которые являются результатом объединения нескольких базовых задач.

Для проверки и контроля усвоения теоретического материала, приобретенных практических навыков работы с технологией CUDA, в течение обучения проводятся мероприятия текущей аттестации в виде контрольной работы и самостоятельных работ на всем протяжении изучения дисциплины. Также проводятся консультации (при необходимости) по разбору заданий для самостоятельной работы, которые вызвали затруднения.

В конце изучения дисциплины студенты сдают зачет. Зачёт выставляется по итогам тестирования и краткого собеседования по его результатам.