

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе и
образовательным инновациям

О. И. Чуприс

« 05 » июля 2018 г.

Регистрационный № УД- 8948 /уч.



ОСНОВЫ ЭФФЕКТИВНОГО ВЫЧИСЛЕНИЯ С ИСПОЛЬЗОВАНИЕМ GPU

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности второй ступени высшего
образования (магистратуры) с углубленной подготовкой специалиста:**

1-31 81 12 Прикладной компьютерный анализ данных

2018 г.

Учебная программа составлена на основе образовательного стандарта высшего образования ОСВО 1-31 81 12-2015 и учебного плана G-31-251/уч. от 26.05.2017.

СОСТАВИТЕЛЬ:

В.В. Пьянов, ассистент кафедры математического моделирования и анализа данных факультета прикладной математики и информатики Белорусского государственного университета.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой математического моделирования и анализа данных Белорусского государственного университета (протокол № 13 от 29 марта 2018);

Научно-методическим Советом Белорусского государственного университета (протокол № 5 от 4 мая 2018 г.).

Зав. кафедрой ММАН Тима /Богачи И.А./

С.А.А.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цель преподавания учебной дисциплины – ознакомление студентов II ступени (магистрантов) основами эффективного программирования для решения высокопроизводительных задач.

В рамках поставленной цели **задачи** учебной дисциплины состоят в следующем:

- 1) изучение синтаксиса программирования с использованием GPGPU;
- 2) формирование практических умений и навыков работы с утилитами NVIDIA;
- 3) приобретение практических навыков для решения прикладных вычислительных задач.

Учебная дисциплина «Основы эффективного вычисления с использованием GPU» относится к циклу дисциплин специальной подготовки компонента учреждения высшего образования.

Учебная программа составлена с учетом межпредметных связей с учебными дисциплинами. Так, основой для изучения дисциплины «Основы эффективного вычисления с использованием GPU» являются дисциплины первой ступени «Программирование» и «Операционные системы». Знания, полученные в результате изучения дисциплины, будут использованы при изучении дисциплины II ступени высшего образования «Методы машинного обучения», а также способствовать успешному прохождению производственной практики по специальности и подготовки магистерской диссертации.

В результате освоения учебной дисциплины студент магистратуры должен:

знать:

- архитектуру GPU;
- принципы построения эффективного ядра CUDA;
- типы памяти GPU, их отличия;
- средства создания ПО для GPU;

уметь:

- создавать программное обеспечение для GPU;
- производить контроль корректности программ GPU;
- производить идентификацию критичных участков программы;
- проектировать эффективное программное обеспечение для GPU;
- обеспечивать безопасность программы от ошибок синхронизации;

владеть:

- навыками работы с различными видами памяти GPU;
- навыками работы с VisualProfiler;
- навыками работы с CUDANsight;

Освоение учебной дисциплины «Основы эффективного вычисления с использованием GPU» должно обеспечить формирование следующих социально-личностных и профессиональных компетенций:

социально-личностные компетенции:

СЛК-1. Учитывать социальные и нравственно-этические нормы в социально-профессиональной деятельности.

профессиональные компетенции:

ПК-3. Самостоятельно разрабатывать эффективные численные методы и алгоритмы, а также интегрировать их в компьютерные системы анализа данных.

Структура содержания учебной дисциплины включает такие дидактические единицы, как темы (разделы), в соответствии с которыми разрабатываются и реализуются соответствующие лекционные и семинарские занятия. Примерная тематика занятий приведена в информационно-методической части.

Дисциплина изучается в 1 семестре (II ступень). Всего на освоение учебной дисциплины «Основы эффективного вычисления с использованием GPU» отведено 120 часов, в том числе 40 аудиторных часов, из них: лекции – 20 часов, лабораторные занятия – 20.

Трудоемкость учебной дисциплины составляет 3 зачетные единицы.

Форма текущей аттестации – зачет.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Основы NVIDIACUDA

Тема 1.1. Основы программирования для GPU. Введение в программирование на GPGPU. Архитектура GPU. Модели памяти. Модель параллелизма SIMT, её отличия от SIMD. Технология NVIDIA CUDA. Программное ядро CUDA. Библиотека Thrust. Работа с памятью.

Тема 1.2. Использование языка C++ для GPU. Использование языка C++ при программировании с использованием CUDA. Ограничения языка. Шаблоны параллельной коммуникации.

Тема 1.3. Фундаментальные алгоритмы GPU. Свертка (reduce). Сканирование (scan). Гистограмма (histogram). Уплотнение (compact), сегментированное сканирование (segmented scan), сортировка. Практическое применение некоторых алгоритмов.

Тема 1.4. Оптимизация GPU программ. Профилирование программы. Технологии отладки на GPU. VisualProfiler. Nsight. Примеры параллелизации последовательных алгоритмов. Эффективная работа с памятью.

Раздел 2. Технологии GPGPU

Тема 2.1. OpenCL. Стандарт OpenCL. Универсальность технологии OpenCL. Особенности языка. Программирование при помощи OpenCL. Сравнение с OpenGL, с CUDA. Драйвер OpenCL. Общая виртуальная память. WebCL. ATI Stream Technology: AMD FireStream.

Тема 2.2. Эффективное использование OpenGL и DirectX с CUDA. Интерфейс CUDA для работы с OpenGL. Интерфейс CUDA для работы с DirectX.

Тема 2.3. Ускорение обучения нейронных сетей на GPU. Основные понятия машинного обучения с использованием GPU. Программные библиотеки для работы с нейронными сетями. cuDNN. TensorRT. DeepStream.

Тема 2.4. Использование программных библиотек с ускорением работы на GPGPU. Математические библиотеки, библиотеки для линейной алгебры. cuBLAS. CUDAMathlibrary. cuSPARSE. cuRAND. cuSOLVER. AmgX. Обработка сигналов и мультимедиа на GPU. cuFFT, NVIDIA Performance Primitives, NVIDIA Codec SDK.

Тема 2.5. Обработка видеопотока на GPU. Основы работы с видеопотоком на GPU. Библиотека с открытым исходным кодом OpenCV. FFmpeg. NVIDIA Codec SDK.

Тема 2.6. Методы построения эффективного гетерогенного программного обеспечения. Технологии IntelXeon. Стандарт OpenACC для HPC. Основные директивы компилятора для работы с OpenACC. Система параллелизма SIMD. Сравнение с OpenMP.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Но ме р ра зд ел а, те мы	Название раздела, темы	Количество аудиторных часов			Форма контроля знаний
		Лек ции	Семи нарск ие занят ия	Лабо ра торные заняти я	
1	Основы NVIDIA CUDA	8		8	
1.1	Основы программирования для GPU.	2		2	Устный опрос
1.2	Использования языка C++ для GPU	2		2	Устный опрос
1.3	Фундаментальные алгоритмы GPU	2		2	Опрос. Проверка выполнения лабораторных работ
1.4	Оптимизация GPU программ	2		2	Контрольная работа №1
2	Технологии GPGPU	12		12	
2.1	OpenCL	2		2	Опрос. Проверка выполнения лабораторных работ
2.2	Эффективное использование OpenGL и DirectX с CUDA	2		2	Опрос. Проверка выполнения лабораторных работ
2.3	Ускорение обучения нейронных сетей на GPU	2		2	Опрос. Проверка выполнения лабораторных работ
2.4	Использование программных библиотек с ускорением работы на GPGPU	2		2	Опрос. Проверка выполнения лабораторных работ
2.5	Обработка видеопотока на GPU	2		2	Коллоквиум
2.6	Методы построения эффективного гетерогенного программного обеспечения	2		2	Контрольная работа №2
ИТОГО		20		20	

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. CUDA C programming guide. – 2010.
2. Sanders, J. CUDA by Example: An Introduction to General-Purpose GPU Programming, Portable Documents /J. Sanders, E. Kandrot. – Addison-Wesley Professional, 2010.
3. Stone, J. E. OpenCL: A parallel programming standard for heterogeneous computing systems / J. E. Stone, D. Gohara, G. Shi. //Computing in science & engineering. – 2010. – V. 12. – №. 3. – P. 66–73.
4. Котов, В. М. Алгоритмы и структуры данных: уч. пособие / В. М. Котов, Е. П. Соболевская, А. А. Толстиков. – Минск : БГУ, 2011.

Перечень дополнительной литературы

1. OpenACC—first experiences with real-world applications / S. Wienke [et al.] //Euro-Par 2012 Parallel Processing. – 2012. – P. 859-870.
2. OpenGL programming guide: the official guide to learning OpenGL, version 1.2 / M. Woo[et al.]. – Addison-Wesley Longman Publishing Co., Inc. – 1999.
3. Jang, H. Neural network implementation using cuda and openmp / H. Jang, A. Park, K. Jung K. //Computing: Techniques and Applications, 2008. DICTA'08. DigitalImage. – IEEE, 2008. – P. 155– 161.

Примерный перечень тем для коллоквиумов

1. Обработка видеопотока на GPU.

Рекомендуемая тематика контрольных работ

- 1) Контрольная работа №1. *Оптимизация GPU программ.*
- 2) Контрольная работа №2. *Методы построения эффективного гетерогенного программного обеспечения.*

Методические рекомендации по организации самостоятельной работы обучающихся

Для организации самостоятельной работы студентов магистратуры по учебной дисциплине следует использовать современные информационные технологии: разместить в сетевом доступе комплекс учебных и учебно-методических материалов (учебно-программные материалы, ссылки на учебные издания для теоретического изучения дисциплины, методические указания к лабораторным занятиям, материалы текущего контроля и текущей аттестации, позволяющие определить соответствие учебной деятельности

обучающихся требованиям образовательных стандартов высшего образования и учебно-программной документации, в т.ч. вопросы для подготовки к зачету, задания, тесты, вопросы для самоконтроля, тематика рефератов и др., список рекомендуемой литературы, информационных ресурсов и др.). Эффективность самостоятельной работы студентов магистратуры проверяется в ходе текущего и итогового контроля знаний. Для общей оценки качества усвоения студентами магистратуры учебного материала рекомендуется использование рейтинговой системы.

Перечень рекомендуемых средств диагностики

Для текущего контроля качества усвоения знаний студентами магистратуры используется следующий диагностический инструментарий:

1. Устная форма: устные опросы; защиты отчетов по домашним заданиям, при выполнении студентами магистратуры лабораторных работ; проведение коллоквиума; защита подготовленного студентом магистратуры реферата (рефераты используются для обобщения и систематизации учебного материала; в процессе подготовки реферата студент магистратуры мобилизует и актуализирует имеющиеся умения, приобретает самостоятельно новые знания, необходимые для раскрытия темы, сопоставляя разные позиции и точки зрения).

2. Письменная форма: письменные контрольные работы по отдельным темам учебной дисциплины.

Методика формирования итоговой оценки

Формой текущей аттестации по учебной дисциплине «Основы эффективного вычисления с использованием GPU» учебным планом предусмотрен зачет.

При оценивании реферата внимание обращается на:

- содержание, корректность и последовательность изложения – 35%;
- релевантность и полноту раскрытия темы – 20 %;
- самостоятельность суждений – 35%;
- оформление – 10%.

Рекомендуется использовать рейтинговую оценку знаний студента магистратуры, дающую возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине. Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний в рейтинговую оценку:

- подготовка реферата – 15 %;
- работа на лабораторных занятиях – 35 %;

- контрольные работы – 30 %;
- коллоквиум – 20 %.

Итоговая оценка формируется на основе:

- 1) Правил проведения аттестации студентов (Постановление Министерства образования Республики Беларусь № 53 от 29 мая 2012г.);
- 2) Положение о рейтинговой системе оценки знаний по дисциплине в БГУ (Приказ ректора БГУ от 18.08.2015 № 382-ОД);
- 3) Критериев оценки знаний студентов (письмо Министерства образования от 22.12.2003).

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Методы машинного обучения	Математического моделирования и анализа данных	нет	Оставить содержание учебной дисциплины без изменения, протокол № 13 от 29.03.2018 г.

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО
ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 20__ г.)

Заведующий кафедрой

УТВЕРЖДАЮ

Декан факультета
