

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт физики



УТВЕРЖДАЮ
Проректор по образовательной деятельности КФУ
Проф. Д.А. Гаюровский



01 » июня 2021 г.

подписано электронно-цифровой подписью

Программа дисциплины

Многопоточная обработка больших данных

Направление подготовки: 03.04.03 - Радиофизика

Профиль подготовки: Распределенные интеллектуальные системы

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2021

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и): профессор, д.н. (профессор) Хуторова О.Г. (Кафедра радиоастрономии, Высшая школа киберфизических систем и прикладной электроники), Olga.Khutorova@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-3	Способен применять современные информационные технологии, использовать компьютерные сети и программные продукты для решения задач профессиональной деятельности.
УК-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

знать: возможности высокопроизводительных вычислительных систем, принципы их работы; постановку проблем моделирования сложных систем

Должен уметь:

использовать высокопроизводительные вычислительные системы для решения задач профессиональной деятельности; применять методы анализа больших данных на практике

Должен владеть:

навыками реализации методов параллельного программирования при решении задач радиофизики; методами анализа экспериментальных и модельных данных, используя современные цифровые технологии

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.О.06 Дисциплины (модули)" основной профессиональной образовательной программы 03.04.03 "Радиофизика (Распределенные интеллектуальные системы)" и относится к обязательной части ОПОП ВО.

Осваивается на 1 курсе в 2 семестре.

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 5 зачетных(ые) единиц(ы) на 180 часа(ов).

Контактная работа - 24 часа(ов), в том числе лекции - 12 часа(ов), практические занятия - 12 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 156 часа(ов).

Контроль (зачёт / экзамен) - 0 часа(ов).

Форма промежуточного контроля дисциплины: зачет во 2 семестре.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
N	Разделы дисциплины / модуля	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
1.	Тема 1. Введение. Большие данные в задачах радиофизики.	2	2	0	0	0	0	0	8
2.	Тема 2. Архитектура современных многопроцессорных вычислительных систем.	2	4	0	0	0	0	0	8
3.	Тема 3. Программная организация высокопроизводительных вычислительных систем.	2	6	0	6	0	0	0	70
4.2. Содержание дисциплины (модуля)									
	Тема 4. Методы анализа больших данных.			0	6	0	0	0	70
Тема 1. Введение. Большие данные в задачах радиофизики.									
Введение. Высокопроизводительные вычислительные системы и области их применения. Производительность вычислительных систем. Цели и назначение курса									
			12	0	12	0	0	0	156
Основные разделы курса. Методические рекомендации по изучению курса									

Обзор литературы. Краткая история компьютерных технологий. Производительность суперкомпьютеров. TOP 500.

Тема 2. Архитектура современных многопроцессорных вычислительных систем.

Архитектура современных многопроцессорных вычислительных систем. Особенности архитектуры для увеличения производительности компьютеров. Классификации вычислительных систем Флинна. Классификации вычислительных систем по типу распределения памяти. Распределенные системы и грид-технологии. Вычисления на видеокартах. Технология Cuda/

Тема 3. Программная организация высокопроизводительных вычислительных систем.

Особенности программирования параллельных вычислений. Модели вычислений и ресурсы параллелизма. Показатели эффективности параллельного алгоритма. Законы Амдала

Этапы разработки параллельного алгоритма. Программный уровень организации параллельных вычислений.

Декомпозиция (тривиальная, функциональная, по данным). Планирование коммуникаций.

Практика программирования с использованием современных библиотек на примере OpenMP. Модель программирования и Терминология OpenMP. Директивы и Основные функции OpenMP.

Директивы препроцессора. Функции OpenMP. Переменные окружения. Общая и разделяемая память.

Многопоточность. Работа с циклами. Синхронизация потоков. Оценка временной сложности.

Тема 4. Методы анализа больших данных.

Параллельная и многопоточная работа в Python. Использование Google Colab для многопоточной обработки данных.

Практика программирования с использованием современных библиотек на примере PyCuda. Методы анализа больших данных. Майнинг данных

Работа с данными в Python. NumPy и Pandas. Понижение размерности и регрессия. Фурье анализ. Вейвлет анализ.

Классификация данных. Кластерный анализ. Методы библиотек SciPy, OpenCV, Scikit-Learn.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства науки и высшего образования Российской Федерации от 6 апреля 2021 года №245)

Письмо Министерства образования Российской Федерации №14-55-99бин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

Google Colab - <https://colab.research.google.com/>

Рейтинг суперкомпьютеров - TOP500 - <http://www.top500.org/>

сайт о параллельных технологиях - <http://www.parallel.ru/>

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;

- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Experimental Parallel Algorithmics - <http://www.umiacs.umd.edu/research/EXPAR/>

Лекция кластеры и MPI - <https://www.youtube.com/watch?v=LNw6dbCQrEE>

Рейтинг суперкомпьютеров СНГ - <http://top50.supercomputers.ru/>

Рейтинг суперкомпьютеров TOP500 - <http://www.top500.org/>

сайт о параллельных технологиях - <http://www.parallel.ru/>

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Вид работ	Методические рекомендации
лекции	Тип занятия: ознакомление с новым материалом. Форма занятия: лекция в аудитории. Перед лекцией рекомендуется предварительно ознакомиться с программой курса, рекомендованными источниками, конспектирование рекомендованных источников с целью выделения основной цели лекции. На лекции рекомендуется краткий конспект основных моментов, активная работа и обсуждение спорных вопросов с преподавателем.
практические занятия	Тип занятия: практическая (лабораторная) работа или семинар. Форма занятия: решение задач на компьютере (самостоятельно и в аудитории под руководством преподавателя). Цель - научить применять полученные знания на практике, выявить готовность и способность использовать в своей научно-исследовательской деятельности, осуществить контроль обучения, продолжить систематизацию знаний.
самостоятельная работа	Основными видами самостоятельной работы являются: 1) предварительная подготовка к аудиторным занятиям. Такая подготовка предполагает изучение учебной программы, установление связи с ранее полученными знаниями, выделение наиболее значимых и актуальных проблем, на изучении которых следует обратить особое внимание и др.; 2) самостоятельная работа при осмыслении учебной информации, сообщаемой преподавателем, ее обобщение и краткая запись, а также своевременная доработка конспектов лекций; 3) подбор, изучение, анализ и при необходимости - конспектирование рекомендованных источников по учебным дисциплинам; 4) выяснение наиболее сложных, непонятных вопросов и их уточнение во время консультаций; 5) подготовка к контрольным занятиям, зачетам; 6) выполнение специальных учебных заданий, предусмотренных учебной программой;
зачет	При подготовке к зачету студент должен правильно и рационально распланировать свое время, чтобы успеть качественно и на высоком уровне подготовиться к ответам по всем вопросам. Во время подготовки к зачету студенты также систематизируют знания, которые они приобрели при изучении разделов курса. При подготовке к зачету следует изучить предлагаемую литературу, конспекты лекций, ознакомиться с вопросами. Важно помнить, что рекомендуемые учебники и специальная литература при изучении курса, имеются в рекомендованном списке литературы в рабочей программе по данному курсу. Студент в целях получения качественных и системных знаний должен начинать подготовку к зачету заранее. Для этого, как уже отмечалось, имеются в программе дисциплины примерные вопросы к зачету. Целесообразно при изучении курса пользоваться его учебно-методическим комплексом. На зачете Студенту предлагается ответить на 2 вопроса по выбранному билету, на подготовку к которым отводится 40 минут. На каждый вопрос студент отвечает 5-15 минут, еще 5 минут отводится на дополнительный вопрос, который может быть задан преподавателем из любого раздела курса по списку вопросов к зачету, выданных Студентам.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Мультимедийная аудитория.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 03.04.03 "Радиофизика" и магистерской программе "Распределенные интеллектуальные системы".

*Приложение 2
к рабочей программе дисциплины (модуля)
Б1.О.06 Многопоточная обработка больших данных*

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 03.04.03 - Радиофизика
Профиль подготовки: Распределенные интеллектуальные системы
Квалификация выпускника: магистр
Форма обучения: очное
Язык обучения: русский
Год начала обучения по образовательной программе: 2021

Основная литература:

1. Параллельные вычисления общего назначения на графических процессорах: учебное пособие / Некрасов К.А., Поташников С.И., Боярченко А.С., - 2-е изд., стер. - Москва: Флинта, Изд-во Урал. ун-та, 2017. - 102 с. ISBN 978-5-9765-3182-6. - Текст: электронный. - URL: <https://znanium.com/catalog/product/949717> (дата обращения: 19.05.2021). - Режим доступа: по подписке.
2. Федотов, И. Е. Параллельное программирование. Модели и приемы : практическое пособие / И. Е. Федотов. - Москва : СОЛОН-Пресс, 2020. - 390 с. - (Серия 'Библиотека профессионала'). - ISBN 978-5-91359-222-4. - Текст : электронный. -

Дополнительная литература:

1. Кузьмич, Р.И. Вычислительные системы, сети и телекоммуникации: учебное пособие / Р.И. Кузьмич, А.Н. Пупков, Л.Н. Корпачева. - Красноярск: Сиб. федер. ун-т, 2018. - 120 с. - Текст: электронный. - Текст: электронный. - URL: <http://znanium.com/catalog/product/1032192> (дата обращения: 21.05.2021). -- Режим доступа: по подписке.
2. Карпова, Е. Д. Основы многопоточного и параллельного программирования: учебное пособие / Карпова Е.Д. - Красноярск: СФУ, 2016. - 356 с.: ISBN 978-5-7638-3385-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/966962> (дата обращения: 21.05.2021). -- Режим доступа: по подписке.
3. Тептин Г. М. Введение в современные высокопроизводительные вычислительные системы: учебно-методическое пособие / Тептин Г. М., Хуторова О. Г., Зинин Д. П. ФГАОУ ВПО 'Казан. (Приволж.) федер. ун-т', Ин-т физики, Каф. радиоастрономии. - [Изд. 2-е, испр.]. - Электронные данные (1 файл: 0,73 Мб)0. - (Казань : Казанский федеральный университет, 2015). - Загл. с экрана.- Казань, 2015. - Текст: электронный. - URL: http://libweb.kpfu.ru/ebooks/06-IPh/06_45_001012.pdf (дата обращения: 21.05.2021). -- Режим доступа: открытый.
4. Маккинли, У. Python и анализ данных / Уэс Маккинли - Москва : ДМК Пресс, 2015. - 482 с. - ISBN 978-5-97060-315-4. - Текст : электронный // ЭБС 'Консультант студента' : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970603154.html> (дата обращения: 21.05.2021). - Режим доступа: по подписке.

*Приложение 3
к рабочей программе дисциплины (модуля)
Б1.О.06 Многопоточная обработка больших данных*

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 03.04.03 - Радиофизика

Профиль подготовки: Распределенные интеллектуальные системы

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2021

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "Консультант студента", доступ к которой предоставлен обучающимся. Многопрофильный образовательный ресурс "Консультант студента" является электронной библиотечной системой (ЭБС), предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Полностью соответствует требованиям федеральных государственных образовательных стандартов высшего образования к комплектованию библиотек, в том числе электронных, в части формирования фондов основной и дополнительной литературы.