

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»

УТВЕРЖДАЮ:

Директор ИЭФ

Ю.И. Соколов

29 мая 2020 г.

Кафедра «Информационные системы цифровой экономики»

Авторы Медникова Оксана Васильевна, к.т.н.

Биленко Александр Васильевич

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Методы и системы обработки больших данных

Направление подготовки:	09.03.03 – Прикладная информатика
Профиль:	Прикладная информатика в бизнесе
Квалификация выпускника:	Бакалавр
Форма обучения:	очная
Год начала подготовки	2020

Одобрено на заседании Учебно-методической комиссии института Протокол № 6 20 мая 2020 г. Председатель учебно-методической комиссии М.В. Ишханян	Одобрено на заседании кафедры Протокол № 15 12 мая 2020 г. Заведующий кафедрой Л.А. Каргина
---	---

Рабочая программа учебной дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 564169
Подписал: Заведующий кафедрой Каргина Лариса Андреевна
Дата: 12.05.2020

Москва 2020 г.

1. ЦЕЛИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Цель– дисциплина предназначена для того, чтобы дать знания, умения и основные навыки, позволяющие создавать высокопроизводительные реализации известных методов вычислительной математики, анализа и обработки данных. Целью освоения дисциплины является – освоение базовых знаний в области архитектуры современных многопроцессорных вычислительных систем параллельной обработки информации, технологий организации параллельных вычислений на многопроцессорных вычислительных комплексах с распределенной или общей оперативной памятью. В результате освоения дисциплины студент должен:

знать:

- основные направления развития высокопроизводительных компьютеров;
- основные классификации многопроцессорных вычислительных систем;
- основные подходы к разработке параллельных программ;
- основные технологии и модели параллельного программирования;
- методы параллельных вычислений для задач вычислительной математики (матричные вычисления, решение систем линейных уравнений, сортировка, обработка графов, уравнения в частных производных, многоэкстремальная оптимизация);

уметь:

- создавать параллельные программы для вычислительных систем с распределенной, общей оперативной памятью;
- проводить распараллеливание вычислительных алгоритмов;
- строить модель выполнения параллельных программ;
- оценивать эффективности параллельных вычислений;
- анализировать сложность вычислений и возможность распараллеливания разрабатываемых алгоритмов;
- применять общие схемы разработки параллельных программ для реализаций собственных алгоритмов;
- оценивать основные параметры получаемых параллельных программ, таких как ускорение, эффективность и масштабируемость;

иметь навыки:

- создания параллельные программы для вычислительных систем с распределенной, общей оперативной памятью;
- построения параллельных аналогов вычислительных алгоритмов.

2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Учебная дисциплина "Методы и системы обработки больших данных" относится к блоку 1 "Дисциплины (модули)" и входит в его вариативную часть.

2.1. Наименования предшествующих дисциплин

Для изучения данной дисциплины необходимы следующие знания, умения и навыки, формируемые предшествующими дисциплинами:

2.1.1. Объектно-ориентированный анализ и программирование:

Знания: имеет сведения об основных языках программирования и работы с базами данных, операционными системами и оболочками, современных программных средах разработки информационных систем и технологий

Умения: активно применяет языки программирования и работы с базами данных, современные программные среды разработки информационных систем и технологий для автоматизации бизнес-процессов, решения прикладных задач различных классов, ведения баз данных и информационных хранилищ

Навыки: программирует, отлаживает и тестирует прототипы программно-технических комплексов задач

2.2. Наименование последующих дисциплин

Результаты освоения дисциплины используются при изучении последующих учебных дисциплин:

2.2.1. Математические методы прогнозирования экономических процессов в бизнесе

Знания: основные принципы аналитического представления и математические законы, позволяющие их обрабатывать

Умения: интерпретировать состояния и действия объектов с помощью математических представлений

Навыки: аналитическими методами синтеза комбинационных схем с заданными параметрами

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫЕ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

В результате освоения дисциплины студент должен:

№ п/п	Код и название компетенции	Ожидаемые результаты
1	ПКС-4 Способен принимать решения по управлению техническими, программно-технологическими и информационно-коммуникационными ресурсами	ПКС-4.1 Управление эффективностью ресурсного обеспечения ИТ проекта. ПКС-4.2 Управление взаимоотношениями с заинтересованными лицами в процессе обеспечения техническими, технологическими информационно-коммуникационными ресурсами.

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ И АКАДЕМИЧЕСКИХ ЧАСАХ

4.1. Общая трудоемкость дисциплины составляет:

4 зачетных единиц (144 ак. ч.).

4.2. Распределение объема учебной дисциплины на контактную работу с преподавателем и самостоятельную работу обучающихся

Вид учебной работы	Количество часов	
	Всего по учебному плану	Семестр 6
Контактная работа	76	76,15
Аудиторные занятия (всего):	76	76
В том числе:		
лекции (Л)	38	38
практические (ПЗ) и семинарские (С)	38	38
Самостоятельная работа (всего)	32	32
Экзамен (при наличии)	36	36
ОБЩАЯ трудоемкость дисциплины, часы:	144	144
ОБЩАЯ трудоемкость дисциплины, зач.ед.:	4.0	4.0
Текущий контроль успеваемости (количество и вид текущего контроля)	ПК1, ПК2	ПК1, ПК2
Виды промежуточной аттестации (экзамен, зачет)	ЭК	ЭК

4.3. Содержание дисциплины (модуля), структурированное по темам (разделам)

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
1	6	Раздел 1 Введение в понятия высокопроизводительных вычислений. Основные направления развития высокопроизводительных компьютеров.	4				18	22	
2	6	Тема 1.1 Важность проблематики параллельных вычислений Пути достижения параллелизма. Векторная и конвейерная обработка данных. Многопроцессорная и многомашинная, параллельная обработка данных. Закон Мура, сдерживающие факторы наращивания количества транзисторов на кристалле и частоты процессоров. Привлекательность подхода параллельной обработки данных.	2				9	11	
3	6	Тема 1.2 Сдерживающие факторы повсеместного внедрения параллельных вычислений Ведомственные, национальные и другие программы, направленные на развитие параллельных вычислений в России. Необходимость изучения дисциплины параллельного программирования. Перечень критических задач, решение которых без использования параллельных вычислений затруднено или вовсе невозможно.	2				9	11	
4	6	Раздел 2 Классификация	28		10		14	88	

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
		многопроцессорных вычислительных систем.							
5	6	Тема 2.1 Системы с распределенной, общей памятью, примеры систем. Массивно-параллельные системы (MPP). Симметричные мультипроцессорные системы (SMP). Параллельные векторные системы (PVP).	20		10		14	80	ПК1, ПК2, ЭК
6	6	Тема 2.2 Системы с неоднородным доступом к памяти (Numa) Компьютерные кластеры – специализированные и полнофункциональные. История возникновения компьютерных кластеров–проект Beowulf. Ме-та-компьютинг. Классификация Флинна, Шора и т.д. Организация межпроцессорных связей – коммуникационные топологии. Примеры сетевых решений для создания кластерных систем	8					8	ПК1, ситуационные задачи
7	6	Раздел 3 Основные принципы организации параллельной обработки данных: модели, методы и технологии параллельного программирования	4		10			14	
8	6	Тема 3.1 Функциональный параллелизм, параллелизм по данным. Парадигма master-slave. Парадигма SPMD. Парадигма конвейеризации. Парадигма «разделяй и властвуй». Спекулятивный параллелизм. Важность	2		6			8	

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
		выбора технологии для реализации алго-ритма							
9	6	Тема 3.2 Модель обмена сообщениями – MPI. Модель общей памяти – OpenMP. Концепция виртуальной, разделяемой памяти – Linda. Российские разработки – Т-система, система DVM. Проблемы создания средства автоматического распараллеливания программ	2		4			6	
10	6	Раздел 4 Параллельное программирование с использованием интерфейса передачи сообщений MPI.	2		4			6	
11	6	Тема 4.1 Библиотека MPI Библиотека MPI. Модель SIMD. Инициализация и завершение MPI-приложения. Точечные обмены данными между процессами MPI-программы. Режимы буферизации. Проблема deadlock'ов. Коллективные взаимодействия процессов в MPI. Управление группами и коммутаторами в MPI	2		4			6	
12	6	Раздел 5 Параллельное программирование на системах с общей памятью (OpenMP)			8			8	
13	6	Тема 5.1 Введение в OpenMP Стандарты программирования для систем с разделяемой памятью. Создание многопоточных приложений. Использование многопоточности при программировании для многоядерных платформ.			8			8	

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
		Синхронизация данных между ветвями в параллельной программе. Директивы языка OpenMP							
14	6	Раздел 6 Параллельное программирование много-ядерных GPU. Кластеры из GPU и супер-компьютеры на гибридной схеме			6			6	
15	6	Тема 6.1 Существующие многоядерные системы. GPU. Использование OpenMP и MPI технологий совместно с CUDA. Степень параллелизма численного алгоритма. Закон Амдала. Параллельный алгоритм решения СЛАУ			6			6	
16		Раздел 7 зачет							
17		Всего:	38		38		32	144	

4.4. Лабораторные работы / практические занятия

Лабораторные работы учебным планом не предусмотрены.

Практические занятия предусмотрены в объеме 38 ак. ч.

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Наименование занятий	Всего часов/ из них часов в интерактивной форме
1	2	3	4	5
1	6	РАЗДЕЛ 2 Классификация многопроцессорных вычислительных систем.	Системы с распределенной, общей памятью, примеры систем. Массивно-параллельные системы (MPP). Симметричные мультипроцессорные системы (SMP). Параллельные векторные системы (PVP).	10
2	6	РАЗДЕЛ 3 Основные принципы организации параллельной обработки данных: модели, методы и технологии параллельного программирования Тема: Функциональный параллелизм, параллелизм по данным.	Работа на портале ИЭФ ml.miit-ief.ru с контентом по текущей теме	6
3	6	РАЗДЕЛ 3 Основные принципы организации параллельной обработки данных: модели, методы и технологии параллельного программирования Тема: Модель обмена сообщениями – MPI.	Работа на портале ИЭФ ml.miit-ief.ru с контентом по текущей теме	4
4	6	РАЗДЕЛ 4 Параллельное программирование с использованием интерфейса передачи сообщений MPI. Тема: Библиотека MPI	Подготовка, выполнение и защита лабораторных работ [2, с.24-36]	4
5	6	РАЗДЕЛ 5 Параллельное программирование на системах с общей памятью (OpenMP) Тема: Введение в OpenMP	Изучение основной и дополнительной литературы [1, с.12-15]	4

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Наименование занятий	Всего часов/ из них часов в интерактивной форме
1	2	3	4	5
6	6	РАЗДЕЛ 5 Параллельное программирование на системах с общей памятью (OpenMP) Тема: Введение в OpenMP	Изучение основной и дополнительной литературы [1, с.12-15]	4
7	6	РАЗДЕЛ 5 Параллельное программирование на системах с общей памятью (OpenMP)	Введение в OpenMP Стандарты программирования для систем с разделяемой памятью. Создание многопоточных приложений. Использование многопоточности при программировании для многоядерных платформ. Синхронизация данных между ветвями в параллельной программе. Директивы языка OpenMP	4
8	6	РАЗДЕЛ 6 Параллельное программирование много-ядерных GPU. Кластеры из GPU и супер-компьютеры на гибридной схеме Тема: Существующие многоядерные системы.	Подготовка к тестированию [1, с.45-75]	6
ВСЕГО:				42/0

4.5. Примерная тематика курсовых проектов (работ)

Курсовые работы (проекты) не предусмотрены.

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В обучении студентов по данной дисциплине используются:

1. при проведении лекционных занятий:

- вводная;
- лекция-информация;
- проблемная лекция;
- лекция визуализация;

2. для проведения лабораторных занятий:

- проектная технология;
- технология учебного исследования;
- техника «круглый стол»,
- техника «публичная защита»;
- технология обучения в сотрудничестве и в малых группах;
- технология проблемного обучения;
- технологии дистанционного обучения;
- разбор конкретных ситуаций.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Вид самостоятельной работы студента. Перечень учебно-методического обеспечения для самостоятельной работы	Всего часов
1	2	3	4	5
1	6	РАЗДЕЛ 1 Введение в понятия высокопроизводительных вычислений. Основные направления развития высокопроизводительных компьютеров. Тема 1: Важность проблематики параллельных вычислений	Работа на портале ИЭФ ml.miit-ief.ru с контентом по текущей теме	9
2	6	РАЗДЕЛ 1 Введение в понятия высокопроизводительных вычислений. Основные направления развития высокопроизводительных компьютеров. Тема 2: Сдерживающие факторы повсеместного внедрения параллельных вычислений	Подготовка, выполнение и защита лабораторных работ [2, с.24-36]	9
3	6	РАЗДЕЛ 2 Классификация многопроцессорных вычислительных систем.	Системы с распределенной, общей памятью, примеры систем. Массивно-параллельные системы (MPP). Симметричные мультипроцессорные системы (SMP). Параллельные векторные системы (PVP).	14
ВСЕГО:				32

7. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Основная литература

№ п/п	Наименование	Автор (ы)	Год и место издания Место доступа	Используется при изучении разделов, номера страниц
1	Высокопроизводительные вычисления: учеб. пособие	Баденко В. Л.	СПб.: Изд-во Политехн. ун-та, 2013 http://www.fb2portal.ru/kostin-s-p/samouchitel-sozdaniya-web-saytov/	1-6 8-239

7.2. Дополнительная литература

№ п/п	Наименование	Автор (ы)	Год и место издания Место доступа	Используется при изучении разделов, номера страниц
2	Параллельное программирование с использованием технологии OpenMP: Учебное пособие	Антонов А.С.	М.: Изд-во МГУ, 2013 http://booksee.org/book/718548	1-6 5-321

8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫЕ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1. <http://smarty.php.net>
2. <http://yandex.ru/cgi/vesna>
3. <http://www.virtech.ru>
4. <http://www.oracle.sun.com>

9. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

1. ОС Windows X.X;
2. Microsoft Office;
3. ASP.NET;
4. Silverlight 4;

10. ОПИСАНИЕ МАТЕРИАЛЬНО ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

1. компьютеры,
2. проектор,

11. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Основными видами аудиторной работы студентов являются лекции и лабораторные работы.

В ходе лекций преподаватель излагает и разъясняет основные, наиболее сложные понятия темы, а также связанные с ней теоретические и практические проблемы, дает рекомендации на лабораторную работу и указания на самостоятельную работу.

Лабораторные работы завершают изучение наиболее важных тем учебной дисциплины.

Они служат для закрепления изученного материала, развития умений и навыков в практической работе по дисциплине: расчет сложности алгоритмов, разработки алгоритмов и программ, подготовки докладов, сообщений, аргументации и защиты лабораторных работ, а также для контроля преподавателем степени подготовленности студентов по изучаемой дисциплине.

Лабораторная работа начинается со вступительного слова преподавателя, формулирующего цель занятия и характеризующего его основную проблематику. Затем, как правило, заслушиваются сообщения студентов. Обсуждение сообщения совмещается с рассмотрением намеченных вопросов на практике. Поощряется выдвижение и обсуждение альтернативных мнений при выполнении практической части лабораторной работы. В заключительном слове преподаватель подводит итоги лабораторной работы и объявляет оценки студентам. В целях контроля подготовленности студентов и привития им навыков практики по дисциплине преподаватель в ходе лабораторной работы может осуществлять текущий контроль знаний в виде тестовых заданий.

При подготовке к лабораторной работе студенты имеют возможность воспользоваться консультациями преподавателя. Кроме указанных тем студенты вправе, по согласованию с преподавателем, избирать и другие интересующие их темы.