

МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

А.В. Замятин

« 02 » сентября 2021 г.



Технологии высокопроизводительной обработки данных
рабочая программа дисциплины

Закреплена за кафедрой Учебный план	<i>Теоретических основ информатики 10.05.01 Компьютерная безопасность, профиль «Анализ безопасности компьютерных систем»</i>
Форма обучения	<i>очная</i>
Общая трудоёмкость	<i>3</i>
Часов по учебному плану	<i>108</i>
в том числе:	
аудиторная контактная работа	<i>54,95</i>
самостоятельная работа	<i>53,05</i>
Вид(ы) контроля в семестрах экзамен	<i>8 семестр – экзамен</i>

Томск-2021

Программу составил
канд. техн. наук,
доцент кафедры теоретических основ информатики



Д.В. Дружинин

Рецензент:
д-р техн. наук, профессор,
профессор кафедры теоретических основ информатики



Ю.Л. Костюк

Рабочая программа дисциплины «Технологии высокопроизводительной обработки данных» разработана в соответствии с образовательным стандартом высшего образования – специалитет, самостоятельно устанавливаемым федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский Томский государственный университет» по специальности 10.05.01 Компьютерная безопасность (Утвержден Ученым советом НИ ТГУ, протокол от 30.06.2021 г. № 06).

Рабочая программа одобрена на заседании кафедры теоретических основ информатики

Протокол от 04 июня 2021 г. № 05

Заведующий кафедрой теоретических основ информатики,
д-р техн. наук, профессор



А.В. Замятин

Рабочая программа одобрена на заседании учебно-методической комиссии института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 17 июня 2021 г. № 05

Председатель УМК ИПМКН,
д-р техн. наук, профессор



С.П. Сущенко

Цель освоения дисциплины

Цель – обучить студентов различным методам организации высокопроизводительных вычислений, сформировать навыки использования различных программных инструментов для организации параллельных вычислений.

1. Место дисциплины в структуре ОПОП

Дисциплина «Технологии высокопроизводительной обработки данных» относится к вариативной части профессионального цикла Блока 1 «Дисциплины (модули)», входит в модуль по выбору «Введение в искусственный интеллект».

Для успешного освоения дисциплины необходимо знание основ информатики, опыт разработки простых программ на одном из языков программирования.

Пререквизиты: «Введение в интеллектуальный анализ данных».

Постреквизиты: нет.

2. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Таблица 1.

Компетенция	Индикатор универсальной компетенции	Код и наименование результатов обучения
ОПК-1 Способен оценивать роль информации, информационных технологий и информационной безопасности в современном обществе, их значение для обеспечения объективных потребностей личности, общества и государства	ИОПК-1.1 Учитывает современные тенденции развития информационных технологий в своей профессиональной деятельности	ОР-1: Знать актуальность, базовую терминологию, архитектуру высокопроизводительных систем. ОР-2: Знать методы, алгоритмы высокопроизводительной обработки данных. ОР-3: Владеть навыками использования углубленных теоретических и практических знаний в области использования технологий высокопроизводительной обработки данных для решения задач интеллектуального анализа данных. ОР-4: Владеть навыками разработки и применения математических методов системного и прикладного программного обеспечения для решения задач научной и проектно-технологической деятельности ОР-5: Уметь применять системное и прикладное программное обеспечение для решения научных и практических задач технологической деятельности
ОПК-2. Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение, в том числе отечественного происхождения, для решения задач профессиональной деятельности	ИОПК-2.3 Формулирует предложения по применению программных средств системного и прикладного назначений, в том числе отечественного производства, используемых для решения задач профессиональной деятельности	
ОПК-3 Способен на основании совокупности математических методов разрабатывать, обосновывать и реализовывать процедуры решения задач профессиональной деятельности	ИОПК-3.2 Осуществляет применение основных понятий, фактов, концепций, принципов математики и информатики для решения задач профессиональной деятельности	
ОПК-14 Способен проектировать базы данных, администрировать	ИОПК-14.1 Понимает модели и структуры данных, физические	

системы управления базами данных в соответствии с требованиями по защите информации	модели баз данных, принципы организации и методы проектирования баз данных, языки и системы программирования баз данных	в области интеллектуального анализа данных. ОР-6: Уметь выбирать методы и средства высокопроизводительной обработки данных для решения научно-исследовательских и прикладных задач. ОР-7: Уметь разрабатывать и применять математические методы для решения научно-исследовательских задач.
ПК-2. Способен осуществлять программирование, тестирование и опытную эксплуатацию ИС с использованием технологических и функциональных стандартов, современных моделей и методов оценки качества и надежности программных средств.	ИПК-2.2 Разрабатывает математические модели, реализуемые в средствах защиты информации; ИПК-2.3 Проводит исследования с целью нахождения наиболее целесообразных практических решений по обеспечению защиты информации	

3. Структура и содержание дисциплины

3.1. Структура и трудоемкость видов учебной работы по дисциплине

Общая трудоемкость дисциплины составляет 3 зачетные единицы, 108 часов.

Таблица 2.

Вид учебной работы	Трудоемкость в академических часах	
	2 семестр	всего
Общая трудоемкость		
Контактная работа:	54.7	54.7
Лекции (Л):	16	16
Практические занятия (ЛР)	32	32
Групповые консультации	4,65	4,65
Промежуточная аттестация	2,3	2,3
Самостоятельная работа обучающегося:	53.05	53.05
- изучение учебного материала	7,35	7,35
- подготовка к практическим занятиям	14	14
- подготовка к рубежному контролю по теме	31,7	31.7
Вид промежуточной аттестации – экзамен	экзамен	экзамен

3.2. Содержание и трудоемкость разделов дисциплины/модуля

Таблица 3.

Код занятия	Наименование разделов и тем и их содержание	Вид учебной работы, занятий, контроля	С е м е с т р	Часы в электронной форме	Всего (час.)	Литература	Код (ы) результата(ов) обучения
	Раздел 1. Введение		8		3	Основная литература: [4], [5] Дополнительная литература: [2]	ОР-1; ОР-2;
1.1.	Актуальность, базовая терминология и тенденции развития.	Лекции			1		
1.2.	Параллелизм компьютерных вычислений.	Лекции			1		
1.3.	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			1		
	Текущий контроль успеваемости	Контрольная работа					
	Раздел 2. Архитектура вычислительных систем		8		5	Основная литература: [4], [5] Дополнительная литература: [2], [5]	ОР-1; ОР-2;
2.1.	Архитектура высокопроизводительных вычислительных систем	Лекции			1		
2.2.	Классификация вычислительных систем	Лекции			1		
2.3.	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			3		
	Раздел 3. Облачные технологии, их свойства и типы				15	Основная литература: [1], [2], [3] Дополнительная литература: [1], [2], [3]	ОР-1; ОР-2; ОР-3; ОР-2.3.1; ОР-4; ОР-6; ОР-6;

						Интернет-ресурсы: [1], [2], [3]	
3.1	Облачные технологии, их свойства и типы	Лекции			3		
		Практические задания			8		
3.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			4		
	Раздел 4. Технология вычислений MapReduce		8		14,6	Основная литература: [1], [3] Интернет-ресурсы: [4]	ОП-1; ОП-2; ОП-3; ОП-4; ОП-5;
4.1	Технология вычислений MapReduce	Лекции			3		
		Практические задания			8		
4.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			3,6		
	Раздел 5. Распределённые файловые системы		8		16	Основная литература: [1], [3] Дополнительная литература: [2] Интернет-ресурсы: [5]	ОП-3; ОП-4; ОП-5; ОП-6; ОП-7;
5.1	Распределённые файловые системы	Лекции			3		
		Практические задания			8		
5.2	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			5		
	Раздел 6. Программирование для высокопроизводительных вычислений.		8		16	Основная литература: [4] Дополнительная литература: [4]	ОП-3; ОП-4; ОП-5; ОП-6; ОП-7;
6.1.	Методы программирования для высокопроизводительных вычислений	Лекции			1		
		Практические задания			4		
6.2.	Методология проектирования параллельных алгоритмов	Лекции			2		
		Практические задания			4		
6.3	Изучение учебного материала, публикаций; подготовка к практическим занятиям	СРС			5		

	Консультации в период теоретического обучения и промежуточной аттестации				4,65		
	Подготовка к промежуточной аттестации в форме экзамена	СРС			31,7		
	Прохождение промежуточной аттестации в форме экзамена	Э	8		2,3		ОП-1; ОП-2; ОП-3; ОП-4; ОП-5; ОП-6; ОП-7

4. Образовательные технологии, учебно-методическое и информационное обеспечение для освоения дисциплины

Текущий контроль освоения дисциплины осуществляется в виде проведения регулярных опросов по пройденному материалу и проверки знаний при допуске к выполнению практических работ.

Также текущий контроль успеваемости осуществляется в виде сдачи практических работ.

Промежуточная аттестация по курсу проводится в форме экзамена.

Самостоятельная работа студента включает:

- изучение разделов курса при подготовке к лекционным и практическим работам;
- самостоятельный поиск необходимых источников информации, включая научно-образовательные ресурсы сети Интернет;
- подготовка ответов на контрольные вопросы;
- подготовка к экзамену.

Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций, и методические материалы, определяющие процедуры оценивания результатов обучения, приведены в Приложении 1 к рабочей программе «Фонд оценочных средств».

Типовые контрольные задания или иные материалы, необходимые для текущей аттестации, и методические материалы, определяющие процедуры оценивания результатов текущей аттестации, приведены в Приложении 2 к рабочей программе «Примерные оценочные средства текущей аттестации».

4.1. Рекомендуемая литература и учебно-методическое обеспечение

№ п/п	Авторы / составители	Заглавие	Издательство	Год издания
Основная литература				
1	S. Srinivasan	Cloud Computing Basics electronic resource	New York, NY : Springer New York : Imprint: Springer,	2014
2	A Ohri	R for Cloud Computing electronic resource : An Approach for Data Scientists	New York, NY : Springer New York : Imprint: Springer	2014
3	Zaigham Mahmood	Cloud Computing: Methods and Practical Approaches	London : Springer London : Imprint: Springer XVII, 347 p. 97 illus.: online resource	2013
4	К. Ю. Богачев	Основы параллельного программирования	Москва БИНОМ. Лаборатория знаний	2015
5	В. П. Гергель	Теория и практика параллельных вычислений : учебное пособие	Москва : Интернет- Университет Информационных Технологий	2016
Дополнительная литература				
1	Christoph Fehling, Frank Leymann, Ralph Retter, Walter Schupeck, Peter Arbitter	Cloud Computing Patterns electronic resource : Fundamentals to Design, Build, and Manage Cloud Applications	Vienna : Springer Vienna : Imprint: Springer	2014

2	Xiaolin Li, Judy Qiu	Cloud Computing for Data-Intensive Applications electronic resource	New York, NY : Springer New York : Imprint: Springer	2014
3	Zaigham Mahmood	Cloud Computing electronic resource : Challenges, Limitations and R&D Solutions	New York, NY : Springer New York : Imprint: Springer	2014
4	А. В. Линеv, Д. К. Боголепов, С. И. Баcтраков	Технологии параллельного программирования для процессоров новых архитектур	Нижегородский гос. ун-т им. Н. И. Лобачевского – Москва : Изд-во Московского университета	2010
Перечень ресурсов информационно-телекоммуникационной сети Интернет				
1		Облачный сервис	http://www.ncloudtech.ru	
2		Статья «Бизнес в облаках. Чем полезны облачные технологии для предпринимателя»	https://kontur.ru/article/s/225	
3	Батура Т.В., Мурзин Ф.А., Семич Д.Ф.	Облачные технологии: основные понятия, задачи и тенденции развития.	http://swsys-web.ru/cloud-computing-basic-concepts-problems.html	
4		MapReduce Tutorial	https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html	
5		HDFS Architecture Guide	https://hadoop.apache.org/docs/r1.2.1/hdfs_desi gn.html	

4.2. Перечень лицензионного и программного обеспечения

Microsoft Visual Studio Community Edition 2013 (C++), библиотека классов, реализующая технологию MapReduce (например, <https://github.com/cdmh/mapreduce>). Все используемые программные продукты являются свободно распространяемыми.

4.3. Оборудование и технические средства обучения

Для материально-технического обеспечения дисциплины требуется наличие компьютерной техники с установленным соответствующим программным обеспечением и другого оборудования, поддерживающего проведение презентаций.

Технические характеристики: Компьютер INTANT i6250_T (i5-6500, 3.2GHz\GA-N110M-S2V, mATX\8ГБ DDR3, 1600МГц\1000ГБ 7200об/мин., 32МБ SATAIII\Microtower 400Вт\DVD±RW\2 x 23,6" АОС I2476VW \кл-ра+мышь). Монитор LCD 23" Philips 234E5QDAB – 13 шт.

Интерактивная доска SMART Board SB480, ультракороткофокусный проектор SMART UF70 с настенным креплением

5. Методические указания обучающимся по освоению дисциплины

Основой обучения является курс лекций, читаемый преподавателем. Важным аспектом овладения перечисленными компетенциями является выполнение практических работ. Для самостоятельной работы и дополнительного расширения круга знаний желательно использовать литературу, приведенную в разделе 4.1.

6. Преподавательский состав, реализующий дисциплину:

Дружинин Денис Вячеславович, канд. техн. наук, доцент кафедры теоретических основ информатики ТГУ.

7. Язык преподавания – русский.