

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

2021 г.

## Технологии высокопроизводительной обработки больших данных

### рабочая программа дисциплины

Закреплена за кафедрой Учебный план	<i>теоретических основ информатики 02.03.03 Математическое обеспечение и администрирование информационных систем, профиль «DevOps-инженерия в администрировании инфраструктуры ИТ-разработки»</i>
Форма обучения	<i>очная</i>
Общая трудоёмкость	<i>3 з.е.</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/зачет/зачет с оценкой	<i>Семестр 8 – экзамен</i>

Программу составил:  
канд. физ.-мат. наук,  
доцент кафедры теоретических основ информатики

Д.В. Дружинин

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

С.П. Сущенко

Рабочая программа дисциплины «Технологии высокопроизводительной обработки больших данных» разработана в соответствии с самостоятельно устанавливаемым образовательным стандартом высшего образования – бакалавриат – федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Томский государственный университет» по направлению подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем (Утвержден Ученым советом НИ ТГУ, протокол от 27.10.2021 г. № 08).

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 больших данных» относится к вариативной части Блока 1 «Дисциплины», входит в модуль «Введение в искусственный интеллект».

Пререквизиты дисциплины: Основы программирования.

Постреквизиты дисциплины: Производственная практика.

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

Таблица 1.

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций)
ПК-3 Способен осуществлять научно-исследовательские и опытно-конструкторские разработки как при исследовании самостоятельных тем, так и разработки по тематике организации ПК	ИПК-3.1 Осуществляет проведение работ по обработке и анализу научно-технической информации и результатов исследований	ОР-3.1.1. Уметь применять системное и прикладное программное обеспечение для решения научных и практических задач технологической деятельности в области интеллектуального анализа данных и биоинформатики. ОР-3.1.2. Уметь выбирать методы и средства высокопроизводительной обработки данных для решения научно-исследовательских и прикладных задач. ОР-3.1.3. Уметь разрабатывать и применять математические методы для решения научно-исследовательских задач.

### 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	
	8 семестр	всего
<b>Общая трудоемкость</b>	108	108
<b>Контактная работа:</b>	54,95	54,95
Лекции (Л):	16	16
Практики (ПЗ)	32	32
Лабораторные работы (ЛР)		
Семинары (СЗ)		
Групповые консультации	2	2
Индивидуальные консультации	2,65	2,65
Промежуточная аттестация	2,3	2,3
<b>Самостоятельная работа обучающегося:</b>	53,05	53,05
- изучение учебного материала	10	10
- подготовка к лабораторным занятиям	11,35	11,35

- подготовка к рубежному контролю по теме/разделу	31,7	31,7
<b>Вид промежуточной аттестации (зачет, зачет с оценкой, экзамен)</b>	<b>Экзамен</b>	<b>Экзамен</b>

### 3.2. Содержание и трудоемкость разделов дисциплины

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
	<b>Раздел 1. Введение</b>		1		<b>6</b>	[4], [5], [2]	ОР-3.1.1, ОР-3.1.2, ОР-3.1.3.
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 лабораторным работам	СРС			4		
	Текущий контроль успеваемости	Контроль ная работа					
	<b>Раздел 2. Архитектура вычислительных систем</b>		1		<b>6</b>	[4], [5], [2], [5]	ОР-3.1.1, ОР-3.1.2, ОР-3.1.3.
2.1.	Архитектура высокопроизводительных вычислительных систем	Лекции			1		
2.2.	Классификация вычислительных систем	Лекции			1		
2.3.	Изучение учебного материала, публикаций; подготовка к лабораторным работам	СРС			4		
	<b>Раздел 3. Облачные технологии, их свойства и типы</b>				<b>18</b>	[1], [2], [3]	ОР-3.1.1, ОР-3.1.2, ОР-3.1.3.
3.1	<b>Облачные технологии, их свойства и типы</b>	Лекции			3		
		Лабораторные работы			8		
3.2	Изучение учебного материала, публикаций; подготовка к лабораторным работам	СРС			7		
	<b>Раздел 4. Технология вычислений MapReduce</b>		1		<b>14</b>	[1], [3]	ОР-3.1.1, ОР-3.1.2, ОР-3.1.3.
4.1	Технология вычислений MapReduce	Лекции			3		
		Лабораторные работы			8		
4.2	Изучение учебного материала, публикаций; подготовка к лабораторным работам	СРС			3		
	<b>Раздел 5. Распределённые файловые системы</b>		1		<b>15</b>	[1], [3]	ОР-3.1.1, ОР-3.1.2, ОР-3.1.3.
5.1	Распределённые файловые системы	Лекции			3		

		Лабораторные работы			8		
5.2	Изучение учебного материала, публикаций; подготовка к лабораторным работам	СРС			4		
	<b>Раздел 6. Программирование для высокопроизводительных вычислений.</b>		1		<b>21</b>	[4]	ОП-3.1.1, ОП-3.1.2, ОП-3.1.3.
6.1.	Методы программирования для высокопроизводительных вычислений	Лекции			1		
		Лабораторные работы			4		
6.2.	Методология проектирования параллельных алгоритмов	Лекции			2		
		Лабораторные работы			7		
6.3	Изучение учебного материала, публикаций; подготовка к лабораторным работам	СРС			4		
	<b>Подготовка к промежуточной аттестации в форме экзамена</b>	СРС			31,7		
	<b>Прохождение промежуточной аттестации в форме экзамена</b>	Э	1		2,3		

#### 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 лабораторных работ.

Также текущий контроль успеваемости осуществляется в виде сдачи лабораторных работ.

Промежуточная аттестация по курсу проводится в форме экзамена.

Самостоятельная работа студента включает:

- изучение разделов курса при подготовке к лекционным и лабораторным заданиям;
- 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, количество страниц
Основная литература				
1.	S. Srinivasan	Cloud Computing Basics electronic resource	New York, NY : Springer New York : Imprint: Springer,	2014 г.
2.	A Ohri	R for Cloud Computing electronic resource : An Approach for Data Scientists	New York, NY : Springer New York : Imprint: Springer	2014 г.
3.	Zaigham Mahmood	Cloud Computing: Methods and Practical Approaches	London : Imprint: Springer	2013 г.
4.	К. Ю. Богачев	Основы параллельного программирования	Москва БИНОМ. Лаборатория знаний	2015 г.
5.	В. П. Гергель	Теория и практика параллельных вычислений: учебное пособие	Москва: Интернет- Университет Информационных Технологий	2016 г.
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 г.
2	Xiaolin Li, Judy Qiu	Cloud Computing for Data- Intensive Applications electronic	New York, NY : Springer New York :	2014 г.

		resource	Imprint: Springer	
3	Zaigham Mahmood	Cloud Computing electronic resource : Challenges, Limitations and R&D Solutions	New York, NY : Springer New York : Imprint: Springer	2014 г.
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 г.

#### **4.2. Базы данных и информационно-справочные системы, в том числе зарубежные**

1. Электронная библиотека (репозиторий) ТГУ [Электронный ресурс] / Электронная библиотека (репозиторий) ТГУ: [сайт]. – [Томск, 2011–2016]. – URL: <http://vital.lib.tsu.ru/vital/access/manager/Index>.
2. Облачный сервис [Электронный ресурс] / URL: <http://www.ncloudtech.ru>

#### **4.3. Перечень лицензионного и программного обеспечения**

MS Windows; MS Office, Microsoft Visual Studio Community Edition.

#### **4.4. Оборудование и технические средства обучения**

Для реализации дисциплины необходимы лекционные аудитории и аудитории для проведения практических занятий. Специальные технические средства (проектор, компьютер и т.д.) требуются для демонстрации материала в рамках изучаемых разделов, проведения защиты проектов в конце семестра. Вся основная и дополнительная литература, необходимая для самостоятельной работы и подготовки к экзамену, имеется в научной библиотеке ТГУ.

#### **5. Методические указания обучающимся по освоению дисциплины**

Основой обучения является курс лекций, читаемый преподавателем. Важным аспектом овладения перечисленными компетенциями является выполнение лабораторных 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 язык.**