

Министерство науки и высшего образования Российской Федерации  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДЕНО:  
Директор  
А. В. Замятин

Рабочая программа дисциплины

**Технологии высокопроизводительной обработки больших данных**

по направлению подготовки

**01.03.02 Прикладная математика и информатика**

Направленность (профиль) подготовки:  
**Математические методы в цифровой экономике**

Форма обучения  
**Очная**

Квалификация  
**Бакалавр**

Год приема  
**2024**

СОГЛАСОВАНО:  
Руководитель ОП  
К.И. Лившиц

Председатель УМК  
С.П. Сущенко

## **1. Цель и планируемые результаты освоения дисциплины**

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-3. Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-3.1. Демонстрирует навыки применения современного математического аппарата для построения адекватных математических моделей реальных процессов, объектов и систем в своей предметной области.

ИОПК-3.2. Демонстрирует умение собирать и обрабатывать статистические, экспериментальные, теоретические и т.п. данные для построения математических моделей, расчетов и конкретных практических выводов.

## **2. Задачи освоения дисциплины**

– Изучить и освоить различные методы организации высокопроизводительных вычислений.

– Сформировать навыки использования различных программных инструментов для организации параллельных вычислений.

## **3. Место дисциплины в структуре образовательной программы**

Дисциплина относится к факультативным дисциплинам, предлагается обучающимся на выбор. Дисциплина входит в факультативный модуль «Введение в искусственный интеллект».

## **4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине**

Восьмой семестр, экзамен.

## **5. Входные требования для освоения дисциплины**

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Введение в интеллектуальный анализ данных».

## **6. Язык реализации**

Русский.

## **7. Объем дисциплины**

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 16 ч.

-практические занятия: 32 ч.

Объем самостоятельной работы студента определен учебным планом.

## **8. Содержание дисциплины, структурированное по темам**

**Тема 1. Актуальность, базовая терминология и тенденции развития. Параллелизм компьютерных вычислений.**

Понятие высокопроизводительных вычислений. Способы повышения производительности вычислительной системы. Понятие параллельных вычислений.

Причины, порождающие вычислительный параллелизм. Мультипроцессирование. Классификация уровней параллелизма, предложенная П. Треливеном. Векторная обработка данных. Многофункциональная обработка данных. Конвейер команд.

**Тема 2. Архитектура высокопроизводительных вычислительных систем.**

Классификация вычислительных систем.

Классификация вычислительных систем Флинна. Системы с одним потоком команд и одним потоком данных. Системы с одним потоком команд и множеством потоков данных. Системы с множеством потоков команд и одним потоком данных. Системы с множеством потоков команд и множеством потоков данных.

Классификация вычислительных систем с множеством потоков команд и множеством потоков данных. Системы с разделяемой памятью. Системы с распределённой памятью. Вычислительные системы с централизованной общей памятью. Вычислительные системы с распределённой общей памятью. Симметричная многопроцессорная архитектура. Архитектура CC-NUMA. Архитектура COMA. Архитектура NCC-NUMA. Вычислители с массовым параллелизмом. Кластеры, их типы.

### **Тема 3. Облачные технологии, их свойства и типы.**

Понятие облачных технологий и вычислений. Свойства облачных технологий. Классификация облачных сервисов по типу ресурса, к которому предоставляется доступ. Классификация облачных сервисов по модели развёртывания.

### **Тема 4. Технология вычислений MapReduce.**

Технология обработки данных MapReduce, её сферы применения. Реализации MapReduce. Пример программы, реализованной с помощью технологии MapReduce.

### **Тема 5. Распределённые файловые системы.**

Понятие распределённых файловых систем. Свойства распределённых файловых систем. Файловые системы Google File System, Hadoop distributed file system, их архитектура, сходства и различия.

### **Тема 6. Программирование для высокопроизводительных вычислений. Методология проектирования параллельных алгоритмов.**

Проблемы параллельного программирования. Методология организации параллельных вычислений для SIMD архитектуры. Методология организации параллельных вычислений для MIMD архитектуры. Показатели качества параллельных методов.

Библиотеки для обмена сообщениями. MPI (Message Passing Interface). OpenMP (Open Multi-Processing). Пример программы с использованием OpenMP.

## **9. Текущий контроль по дисциплине**

Текущий контроль по дисциплине проводится в форме вопросов по лекционному материалу, лабораторных работ и фиксируется в форме контрольной точки не менее одного раза в семестр.

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

## **10. Порядок проведения и критерии оценивания промежуточной аттестации**

Экзамен по окончании изучения курса проводится в письменно-устной форме. Допуском до экзамена является выполнение двух лабораторных работ. При выполнении менее трёх лабораторных работ максимальная оценка на экзамене – 4.

Оценка «отлично» ставится, если студент полноценно ответил на вопрос билета, а также на дополнительный вопрос, требующий аналитического сопоставления знаний, полученных при изучении различных тем данной дисциплины. Оценка «хорошо» ставится, если студент полноценно ответил на вопрос билета, но не ответил на дополнительный вопрос. Оценка «удовлетворительно» ставится, если студент частично ответил на вопрос билета.

Оценочные материалы для проведения промежуточной аттестации размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

## 11. Учебно-методическое обеспечение

- а) Электронный учебный курс по дисциплине в LMS iDo;
- б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.
- в) Учебное пособие.
- г) Видеолекции.
- д) Методические указания по проведению лабораторных работ.

## 12. Перечень учебной литературы и ресурсов сети Интернет

- а) основная литература:
  - S. Srinivasan. Cloud Computing Basics electronic resource. - New York, NY : Springer New York : Imprint: Springer, 2014.
  - A Ohri. R for Cloud Computing electronic resource : An Approach for Data Scientists. - New York, NY : Springer New York : Imprint: Springer, 2014.
  - Zaigham Mahmood. Cloud Computing: Methods and Practical Approaches. - London : Springer London : Imprint: Springer, 2013.
  - Богачев К.Ю. Основы параллельного программирования. - Москва БИНОМ. Лаборатория знаний, 2015.
  - Гергель В.П. Теория и практика параллельных вычислений : учебное пособие. - Москва : Интернет-Университет Информационных Технологий, 2016.

- б) дополнительная литература:
  - 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.
  - Xiaolin Li, Judy Qiu. Cloud Computing for Data-Intensive Applications electronic resource. - New York, NY : Springer New York : Imprint: Springer, 2014.
  - Zaigham Mahmood. Cloud Computing electronic resource : Challenges, Limitations and R&D Solutions. - New York, NY : Springer New York : Imprint: Springer, 2014.
  - А. В. Линева, Д. К. Боголепов, С. И. Бахраков. Технологии параллельного программирования для процессоров новых архитектур. - Нижегородский гос. ун-т им. Н. И. Лобачевского – Москва : Изд-во Московского университет, 2010.

- в) ресурсы сети Интернет:
  - Облачный сервис - <http://www.ncloudtech.ru>
  - Статья «Бизнес в облаках. Чем полезны облачные технологии для предпринимателя» - <https://kontur.ru/articles/225>
  - 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>
  - MapReduce Tutorial - [https://hadoop.apache.org/docs/r1.2.1/mapred\\_tutorial.html](https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html)
  - HDFS Architecture Guide - [https://hadoop.apache.org/docs/r1.2.1/hdfs\\_design.html](https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html)
  - Общероссийская Сеть КонсультантПлюс Справочная правовая система. <http://www.consultant.ru>

## 13. Перечень информационных технологий

- а) лицензионное и свободно распространяемое программное обеспечение:
  - Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office OneNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);
  - публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

- б) информационные справочные системы:
- Электронный каталог Научной библиотеки ТГУ –  
<http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>
  - Электронная библиотека (репозиторий) ТГУ –  
<http://vital.lib.tsu.ru/vital/access/manager/Index>
  - ЭБС Лань – <http://e.lanbook.com/>
  - ЭБС Консультант студента – <http://www.studentlibrary.ru/>
  - Образовательная платформа Юрайт – <https://urait.ru/>
  - ЭБС ZNANIUM.com – <https://znanium.com/>
  - ЭБС IPRbooks – <http://www.iprbookshop.ru/>

#### **14. Материально-техническое обеспечение**

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения практических занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

Для проведения лабораторных работ требуются компьютеры, оснащённые дискретными видеокартами фирмы Nvidia с поддержкой технологии CUDA.

#### **15. Информация о разработчиках**

Дружинин Денис Вячеславович, канд. техн. наук, доцент кафедры теоретических основ информатики института прикладной математики и компьютерных наук НИ ТГУ.