

**Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Московский физико-технический институт  
(национальный исследовательский университет)»**

**УТВЕРЖДЕНО**

**Директор физтех-школы  
прикладной математики и  
информатики**

**А.М. Райгородский**

	<b>Рабочая программа дисциплины (модуля)</b>
<b>по дисциплине:</b>	Параллельные и распределенные вычисления
<b>по направлению:</b>	Прикладная математика и информатика
<b>профиль подготовки:</b>	Прикладная математика, компьютерные науки и инженерия Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
<b>курс:</b>	2
<b>квалификация:</b>	бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 45 час.

Подготовка к экзамену: 30 час.

Всего часов: 135, всего зач. ед.: 3

Количество контрольных работ, заданий: 2

Программу составил: О.Н. Ивченко, заведующий кафедрой

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 10.02.2025

## Аннотация

Курс "Параллельные и распределенные вычисления" посвящен изучению методов и технологий создания высокопроизводительных вычислительных систем, способных эффективно обрабатывать большие объемы данных и решать сложные вычислительные задачи. В рамках курса рассматриваются как теоретические основы параллелизма и распределенности, так и практические аспекты разработки и реализации параллельных и распределенных программ.

### 1. Цели и задачи

#### Цель дисциплины

- познакомить студентов с основами работы многопроцессорных вычислительных систем и дать практический опыт работы с такими системами. Курс состоит из двух модулей, посвященных соответственно параллельным и распределенным системам. В первом модуле рассматриваются системы в практически «идеальных» условиях, где вычислительные узлы и соединения между ними надежные и быстродействующие. Во втором модуле рассматриваются способы построения надежных систем из ненадежных компонент.

#### Задачи дисциплины

- ознакомление с типовыми методами исследования и проектирования параллельных и распределенных вычислений; -приобретение навыков применения современных инструментальных средств и технологий исследования и проектирования параллельных и распределенных вычислений.

### 2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.1 Знает основные правила оформления научных публикаций и научно-технической документации, в том числе с использованием прикладного программного обеспечения
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях

### 3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны знать:

- виды и классификацию многопроцессорных вычислительных систем;
- принципы построения распределенных хранилищ данных;
- принципы построения распределенных экосистем (Hadoop / Spark);
- разницу между (одно-) серверными базами данных и распределенными базами данных;
- модель асинхронных вычислений и связь со степенью изоляций транзакций;
- теорему Фишер-Линч-Патерсон (FLP-теорема);
- знать базовые принципы работы Paxos / Raft;
- знать алгоритмы синхронизации времени (NTP, Cristian's Algorithm)
- стандартные задачи распределенных вычислений (Multicasts, Failure Detectors, Membership, Consensus, RSM).

уметь:

- пользоваться библиотекой для параллельных вычислений OpenMP;
- пользоваться библиотекой для параллельных вычислений MPI;
- пользоваться распределенной файловой системой HDFS;
- пользоваться распределенным фреймворком вычислений Hadoop;
- пользоваться распределенным хранилищем данных Hive;
- уметь пользоваться примитивами распределенных вычислений Lamport Timestamps, Vector Clocks
- решать задачу консенсуса в синхронной системе;
- пользоваться алгоритмами Paxos / Raft.

владеть:

- навыками работы с многопроцессорными вычислительными системами (параллельными и распределенными вычислительными системами, в частности)
- кругозором в выборе архитектурного решения поставленной задачи.

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

##### 4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Введение в параллельные и распределенные вычисления	5	5		5
2	Программирование с использованием MPI (Message Passing Interface)	5	5		8
3	Программирование с использованием OpenMP (Open Multi-Processing)	5	5		9
4	Параллельные алгоритмы	5	5		8
5	Распределенные системы и архитектуры	5	5		7
6	Распределенное программирование и системы обработки больших данных	5	5		8
Итого часов		30	30		45
Подготовка к экзамену		30 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

##### 4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 4 (Весенний)

## 1. Введение в параллельные и распределенные вычисления

Что такое параллелизм и распределенность? Основные понятия: процессы, потоки, параллельные архитектуры (многоядерные процессоры, кластеры, облачные вычисления). Модели параллелизма (PRAM, BSP). Закон Амдала и его значение для масштабируемости. Основные проблемы параллельных вычислений: синхронизация, взаимоблокировки (deadlocks), гонки данных (race conditions). Введение в MPI и OpenMP.

## 2. Программирование с использованием MPI (Message Passing Interface)

Подробное изучение MPI: коммуникативные операции (send, receive, broadcast, gather, scatter), коллективные операции. Разработка и отладка параллельных программ на MPI. Примеры реализации параллельных алгоритмов: сортировка, поиск, матричные операции. Анализ производительности MPI-программ.

## 3. Программирование с использованием OpenMP (Open Multi-Processing)

Подробное изучение OpenMP: директивы компилятора (parallel, for, sections, critical, atomic). Шаринговые переменные и private переменные. Синхронизация в OpenMP. Разработка и отладка параллельных программ на OpenMP. Примеры реализации параллельных алгоритмов (аналогичные MPI). Сравнение OpenMP и MPI.

## 4. Параллельные алгоритмы

Обзор основных параллельных алгоритмов: сортировка (быстрая, слиянием), поиск (глубина, ширина), матричные операции (умножение, разложение LU), алгоритмы графовых задач (поиск в ширину, поиск в глубину). Анализ сложности параллельных алгоритмов. Понятие эффективности и масштабируемости.

## 5. Распределенные системы и архитектуры

Архитектуры распределенных систем (клиент-сервер, peer-to-peer). Протоколы распределенной согласованности (Paxos, Raft). Управление распределенными ресурсами. Понятие консенсуса и согласованности данных. Виртуальные машины и контейнеризация. Облачные вычисления: основные понятия и архитектура.

## 6. Распределенное программирование и системы обработки больших данных

Введение в MapReduce. Hadoop и Spark. Обработка больших данных в распределенных системах. Техники обработки потоковых данных. Проблемы масштабируемости и устойчивости в больших данных. Примеры реализации распределенных приложений. Безопасность в распределенных системах.

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

Компьютер и мультимедийное оборудование (проектор, звуковая система).

## 6. Перечень рекомендуемой литературы

### Основная литература

Гергель В. П. Теория и практика параллельных вычислений / В. П. Гергель. – 2-е изд., испр. – Москва : ИНТУИТ, 2024. – 1 файл. – Электронная версия печатной публикации

### Дополнительная литература

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

Не используются

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

Мультимедийные технологии. MS PowerPoint, демонстрация презентаций.

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

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

Успешное освоение дисциплины требует:

- посещения студентом всех видов аудиторных занятий;
- ведения конспекта в ходе аудиторных занятий;
- качественной самостоятельной подготовки к практическим занятиям, активной работы на них;
- активной самостоятельной и аудиторной работы студента;
- своевременной сдачи преподавателю заданий по аудиторным видам работ.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

<b>по направлению:</b>	Прикладная математика и информатика
<b>профиль подготовки:</b>	Прикладная математика, компьютерные науки и инженерия Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
<b>курс:</b>	<u>2</u>
<b>квалификация:</b>	бакалавр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

**Разработчик:** О.Н. Ивченко, заведующий кафедрой

## 1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.2 Находит, критически анализирует и выбирает информацию, необходимую для решения поставленной задачи
ОПК-2 Способен использовать современные информационные технологии и программные средства при решении задач профессиональной деятельности, соблюдая требования информационной безопасности	ОПК-2.2 Знает и умеет применять численные математические методы и прикладное программное обеспечение для решения научных задач в профессиональной области
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.1 Знает основные правила оформления научных публикаций и научно-технической документации, в том числе с использованием прикладного программного обеспечения
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях

## 2. Показатели оценивания компетенций

В результате изучения дисциплины «Параллельные и распределенные вычисления» обучающийся должен:

### знать:

- виды и классификацию многопроцессорных вычислительных систем;
- принципы построения распределенных хранилищ данных;
- принципы построения распределенных экосистем (Hadoop / Spark);
- разницу между (одно-) серверными базами данных и распределенными базами данных;
- модель асинхронных вычислений и связь со степенью изоляций транзакций;
- теорему Фишер-Линч-Патерсон (FLP-теорема);
- знать базовые принципы работы Paxos / Raft;
- знать алгоритмы синхронизации времени (NTP, Cristian's Algorithm)
- стандартные задачи распределенных вычислений (Multicasts, Failure Detectors, Membership, Consensus, RSM).

### уметь:

- пользоваться библиотекой для параллельных вычислений OpenMP;
- пользоваться библиотекой для параллельных вычислений MPI;
- пользоваться распределенной файловой системой HDFS;
- пользоваться распределенным фреймворком вычислений Hadoop;
- пользоваться распределенным хранилищем данных Hive;
- уметь пользоваться примитивами распределенных вычислений Lamport Timestamps, Vector Clocks
- решать задачу консенсуса в синхронной системе;
- пользоваться алгоритмами Paxos / Raft.

### владеть:

- навыками работы с многопроцессорными вычислительными системами (параллельными и распределенными вычислительными системами, в частности)
- кругозором в выборе архитектурного решения поставленной задачи.

## 3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Опишите параллельный алгоритм сортировки (например, быстрая сортировка, сортировка слиянием).
2. Опишите параллельный алгоритм поиска (например, поиск в ширину, поиск в глубину).
3. Опишите параллельный алгоритм умножения матриц.
4. Как оценить эффективность параллельного алгоритма? Какие метрики используются?
5. В чем разница между алгоритмическим и вычислительным параллелизмом?
6. Опишите архитектуру клиент-сервер. Какие преимущества и недостатки она имеет?
7. Что такое MapReduce? Опишите этапы обработки данных в MapReduce.
8. Опишите архитектуру Hadoop.
9. Что такое консенсус в распределенных системах? Приведите примеры алгоритмов достижения консенсуса.
10. Сравните Hadoop и Spark.

#### **4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся**

1. Инструменты распределенных вычислений
2. Описание распределенной файловой системы HDFS, отказоустойчивость и масштабируемость.
3. Достоинства и недостатки HDFS.
4. Назначение Secondary Namenode в HDFS.
5. Процесс чтения и записи данных в HDFS.
6. Структура кластера Hadoop. Роли серверов. Как обеспечивается отказоустойчивость.
7. Процесс выполнения задачи на кластере. Distributed cache.
8. Формальная модель MapReduce вычислений.
9. Области применения Writable и WritableComparable.
10. Типы компараторов в MapReduce.
11. Какие этапы работы MapReduce-приложения должен реализовать разработчик.
12. MapReduce: Combiner, Comparator, Partitioner.
13. Типы Join в MapReduce.
14. Достоинства и недостатки Hadoop Streaming.
15. Описание системы Hive для работы с большими данными (где хранится метаданные, где хранятся реальные данные, возможности системы).
16. Теория распределенных вычислений
17. Модель вычислений: процессоры и сеть, аналогия с разделяемой памятью. Синхронные/асинхронные вычисления. 18. Разделяемая память/message passing. Синхронизатор. Тезисно объяснить эквивалентность систем без отказов.
19. Типология отказов процессов, отказы сети. Вложенность классов отказов. Определение casual order (happens before) для asynchronous message-passing систем.

#### **Пример билета**

##### **Билет 1:**

1. Определение консенсуса, три его свойства. Альтернативные формулировки консенсуса: Terminating Reliable Broadcast, Weak Interactive Consistency. Их эквивалентность.
2. Теорема Фишера, Линча и Патерсона (FLP) (без доказательства). Решение задачи о консенсусе в синхронной системе. CAP-теорема (с доказательством). Разница между CAP и FLP теоремами.
3. ACID. Степени изоляции транзакций.

#### **Критерии оценивания**

- оценка «отлично (10)» выставляется студенту, показавшему всеохватывающие, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений;



- оценка «отлично (9)» выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений;
- оценка «отлично (8)» выставляется студенту, показавшему всесторонние систематизированные, глубокие знания учебной программы дисциплины и умение применять их на практике при решении конкретных задач, и правильное обоснование принятых решений;
- оценка «хорошо (7)» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (6)» выставляется студенту, если он знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «хорошо (5)» выставляется студенту, если он знает материал, и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности;
- оценка «удовлетворительно (4)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «удовлетворительно (3)» выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он владеет фрагментарно основными разделами учебной программы, необходимыми для дальнейшего обучения и может применять полученные знания по образцу в стандартной ситуации;
- оценка «неудовлетворительно (2)» выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач;
- оценка «неудовлетворительно (1)» выставляется студенту, который не знает формулировок основных понятий дисциплины.

## **5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности**

При проведении экзамена обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося проводится в течение 30 минут.

Во время проведения экзамена обучающиеся могут пользоваться программой дисциплины, а также справочной литературой и вычислительной техникой.