

**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>	Информатика и вычислительная техника
<b>профиль подготовки:</b>	Физтех-школа Прикладной Математики и Информатики кафедра алгоритмов и технологий программирования
<b>курс:</b>	3
<b>квалификация:</b>	бакалавр

Семестр, формы промежуточной аттестации: 6 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 0 час.

лабораторные занятия: 30 час.

Самостоятельная работа: 90 час.

Подготовка к экзамену: 30 час.

Всего часов: 180, всего зач. ед.: 4

Количество контрольных работ, заданий: 2

Программу составили:

А.А. Драль, старший преподаватель

О.Н. Ивченко, старший преподаватель

А.Н. Сальников, канд. физ.-мат. наук, доцент

Я.А. Леонов, ассистент

Р.Г. Липовский, старший преподаватель

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 04.06.2020

## Аннотация

На данном курсе студенты познакомятся с:

- такими подходами к разработке параллельных программ как MPI, OpenMP;
- вычислениями на графических процессорах с помощью CUDA.
- устройством распределённых очередей и хранилищ (ZooKeeper, Kafka, Cassandra);
- основными фреймворками обработки больших данных и их внутренним устройством.

Курс ориентирован на практику. Предполагается программирование как на семинарах, так и в рамках домашних заданий.

## 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

В ходе курса студенты получают практические навыки работы как с параллельными, так и с распределенными системами.

## 2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
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

## 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	Теория распределенных вычислений (уровни изоляции транзакций, CRDT, CAP, FLP, Paxos, Raft)	10		10	30
2	Вычисления на GPU. Технология CUDA	6		6	18
3	Параллельные вычисления на MPI и OpenMP	4		4	12
4	Распределённые вычисления на больших объемах данных (HDFS, MapReduce, Hive, Spark)	10		10	30
Итого часов		30		30	90
Подготовка к экзамену		30 час.			
Общая трудоёмкость		180 час., 4 зач.ед.			

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

Семестр: 6 (Весенний)

1. Теория распределенных вычислений (уровни изоляции транзакций, CRDT, CAP, FLP, Paxos, Raft)

Классификация многопроцессорных вычислительных систем, модели отказов;

Отличие параллельных вычислений и распределенных вычислений;

Базы данных: ACID, уровни изоляции. Распределенные базы данных: CAP, CRDT. AP и CP системы;

Алгоритмы синхронизации времени (NTP, Cristian's Algorithm);

стандартные задачи распределенных вычислений (Multicasts, Failure Detectors, Membership, Consensus, RSM);

Теорема Фишер-Линч-Патерсон (FLP-теорема);

Базовые принципы работы Paxos / Raft;

2. Вычисления на GPU. Технология CUDA

Архитектура графических процессоров.

Device (графический процессор) и Host. Обмен данными между ними.

Оптимизация вычислений на графических процессорах.

3. Параллельные вычисления на MPI и OpenMP

Что такое параллельные вычисления?

Устройство и основные структуры в MPI Система очередей SLURM.

Особенности OpenMP

Использование MPI и OpenMP в рамках одной программы.

#### 4. Распределённые вычисления на больших объемах данных (HDFS, MapReduce, Hive, Spark)

Распределённые файловые системы (GFS, HDFS). Её составляющие. Их достоинства, недостатки и сфера применения. Чтение и запись в HDFS. HDFS APIs: WebUI, shell, Java API

Парадигма MapReduce. Основная идея, формальное описание. Обзор реализаций. API для работы с Hadoop (Native Java API vs. Streaming), примеры

Типы Join'ов и их реализации в парадигме MR. Паттерны проектирования MR (pairs, stripes, составные ключи). PageRank в MR. Планировщик задач в YARN.

SQL поверх BigData. Фреймворк Hive

### 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

Не предусмотрено

Дополнительная литература

Не предусмотрено

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

материалы по параллельным вычислениям: <http://www.parallel.ru/>

Информационно-аналитические материалы по параллельным вычислениям.

<http://www.parallel.ru/>

[http://parallel.ru/tech/tech\\_dev/MPI/mpibook.pdf](http://parallel.ru/tech/tech_dev/MPI/mpibook.pdf)

Документация по MPI. <http://www.mpi-forum.org/>

Документация по OpenMP. <http://www.openmp.org>

Измерение производительности MPI. <http://www.mcs.anl.gov/research/projects/mpi/mpptest/>

<http://parallel.ru/info/parallel/openmp/>

Designing and Building Parallel Programs. <http://www.mcs.anl.gov/~itf/dbpp/>

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

Для практических занятий:

- Компьютерный класс. Каждый компьютер должен иметь выход в интернет и ПО для подключения к удалённым серверам.

- Удалённый кластер с такими характеристиками:

Кол-во машин

Характеристики одной машины

Объём оперативной памяти

Кол-во ядер CPU Объём дисковой памяти

Операционная система

1

8

2

200

Linux Ubuntu 16.04

## Linux Ubuntu 16.04

На кластере должен быть развернута последняя версия Cloudera Manager, в который нужно встроить такие сервисы: HDFS, YARN, Hive, Spark2 on YARN, HBase, Zookeeper, Kafka.

- Аккаунты на научном кластере МФТИ:
  - 1 аккаунт для каждого студента
  - 1 расширенный аккаунт для каждого преподавателя (из аккаунта преподавателя должны быть доступны все аккаунты студентов).

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

Успешное освоение курса требует большой самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя:

- проработку учебного материала (по конспектам лекций, учебной и научной литературе), подготовку ответов на вопросы, предназначенных для самостоятельного изучения, доказательство отдельных утверждений, свойств;
  - подготовку к практическим занятиям, выполнение 6 индивидуальных домашних заданий.
- Промежуточный контроль знаний проводится в виде письменных опросов (мини-тестов) по теории.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

**по направлению:** Информатика и вычислительная техника  
**профиль подготовки:** Физтех-школа Прикладной Математики и Информатики  
кафедра алгоритмов и технологий программирования  
**курс:** 3  
**квалификация:** бакалавр

Семестр, формы промежуточной аттестации: 6 (весенний) - Экзамен

**Разработчики:**

А.А. Драль, старший преподаватель  
О.Н. Ивченко, старший преподаватель  
А.Н. Сальников, канд. физ.-мат. наук, доцент  
Я.А. Леонов, ассистент  
Р.Г. Липовский, старший преподаватель

## 1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
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

## 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

Примеры заданий на контрольную №1.

1. Строгий системный администратор хочет знать, на какие машины кластера (клиент, DataNode, NameNode, SecondaryNameNode) вам нужен доступ для штатной работы с HDFS. Ответ нужно ему обосновать.
2. Чем определяется число мапперов? А число редьюсеров? Можно ли задать эти значения?

Примеры заданий на контрольную №2.

1. Охарактеризуйте Kafka в терминах CAP, объясните свой выбор.
2. Чем отличается HBase от Cassandra в терминах CAP.

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

Инструменты распределенных вычислений

Описание распределенной файловой системы HDFS, отказоустойчивость и масштабируемость.

Достоинства и недостатки HDFS.

Назначение Secondary Namenode в HDFS.

Процесс чтения и записи данных в HDFS.

Структура кластера Hadoop. Роли серверов. Как обеспечивается отказоустойчивость.

Процесс выполнения задачи на кластере. Distributed cache.

Формальная модель MapReduce вычислений.

Области применения Writable и WritableComparable.

Типы компараторов в MapReduce.

Какие этапы работы MapReduce-приложения должен реализовать разработчик.

MapReduce: Combiner, Comparator, Partitioner.

Типы Join в MapReduce.

Достоинства и недостатки Hadoop Streaming.

Описание системы Hive для работы с большими данными (где хранится метаданная, где хранятся реальные данные, возможности системы).

Теория распределенных вычислений

Модель вычислений: процессоры и сеть, аналогия с разделяемой памятью.

Синхронные/асинхронные вычисления. Разделяемая память/message passing. Синхронизатор.

Тезисно объяснить эквивалентность систем без отказов.

Типология отказов процессов, отказы сети. Вложенность классов отказов. Определение casual order (happens before) для asynchronous message-passing систем.

Определение консенсуса, три его свойства. Альтернативные формулировки консенсуса: Terminating Reliable Broadcast, Weak Interactive Consistency. Их эквивалентность.

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.

ACID. Степени изоляции транзакций.

Пример билета:

1. Исходные данные: id пользователя, id сессии, url.

Для каждого пользователя посчитайте число сессий и уникальных доменов. Сессия – это посещения человеком страниц (любых страниц, а не страниц одного сайта) с перерывами не более, чем на 30 минут.

Опишите схему решения для Hadoop.

2. Нам нужно считать ТОП N самых популярных слов в тексте с помощью map-reduce программы для Hadoop. Как это можно эффективно реализовать (взяв за основу стандартную программу WordCount), зная N заранее (N порядка 1000).

#### Критерии оценивания

- оценка «отлично (10)» выставляется студенту, показавшему всесторонние, 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**

Итоговая оценка по курсу складывается из оценки за выполненные в ходе семестра практические задания и оценки за ответы на экзамене. Оценка по каждой составляющей должна быть не менее 3х.

2.1.1. Накопленные баллы за работу в семестре (max: 5 баллов)

2.1.2. Ответ на экзамене (max: 5 баллов)

При проведении экзамена обучающемуся предоставляется 60 минут на подготовку. Опрос обучающегося по билету на экзамене не должен превышать двух астрономических часов. Во время проведения экзамена обучающиеся могут пользоваться программой дисциплины, но НЕ могут пользоваться своими конспектами.

Оценка по 5-бальной шкале усредняется за ответ на билет по распределенной и параллельной части.

Регламент проведения экзамена по распределённой части курса:

Оценка складывается из трех составляющих:

- [2 балла] теоретическая часть: ответ на 2 вопроса (один по модулю “Теория распределенных вычислений”, второй - по модулю “Инструменты распределенных вычислений”)
- [1.5 балла] практическая задача. Задача не предполагает программирования, на выходе от студента достаточно получить описание MapReduce-модели задачи или же общей структуры Hive-запроса.
- [1.5 балла] доп. вопросы.