

Минобрнауки России
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)**

УТВЕРЖДАЮ



Заведующий кафедрой
Кургалин Сергей Дмитриевич
Кафедра цифровых технологий

03.05.2023

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.О.10 Высокопроизводительные вычисления и большие данные

1. Код и наименование направления подготовки/специальности:

02.04.01 Математика и компьютерные науки

2. Профиль подготовки/специализация:

Компьютерное моделирование и искусственный интеллект

3. Квалификация (степень) выпускника:

Магистратура

4. Форма обучения:

Очная

5. Кафедра, отвечающая за реализацию дисциплины:

Кафедра цифровых технологий

6. Составители программы:

Борзунов Сергей Викторович, кандидат физико-математических наук, доцент

7. Рекомендована:

протокол НМС ФКН № 7 от 03.05.2023

8. Учебный год:

2023-2024

Семестры:

2

9. Цели и задачи учебной дисциплины:

Целями освоения учебной дисциплины являются: формирование знаний, умений и навыков в сфере разработки и эксплуатации аппаратного и программного обеспечения современных высокопроизводительных распределенных систем, формирование у обучающихся теоретических знаний и практических навыков по работе с большими данными..

Задачи учебной дисциплины:

- знакомство с архитектурой многопроцессорных вычислительных систем;
- обзор средств распараллеливания программного кода на системах с общей и распределенной памятью;
- получение начальных знаний и умений по созданию систем обработки Big Data и использованию BigData в системах реального времени.

10. Место учебной дисциплины в структуре ООП:

Обязательная часть, блок Б1. Для успешного освоения дисциплины необходимо предварительное

изучение следующих разделов математики и компьютерных наук: математический анализ, программирование.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников) и индикаторами их достижения:

Код и название компетенции	Код и название индикатора компетенции	Знания, умения, навыки
ОПК-2 Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.1 Владеет навыками создания и исследования новых математических моделей в естественных науках	Знает области применения высокопроизводительных вычислений
ОПК-2 Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.2 Умеет использовать их в профессиональной деятельности	Умеет использовать технологии высокопроизводительных вычислений для создания и исследования новых математических моделей в естественных науках
ОПК-2 Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.3 Имеет практический опыт создания и исследования подобных математических моделей и разработки теорий и методов для их описания	Владеет навыками создания и исследования математических моделей с использованием технологий высокопроизводительных вычислений
ОПК-3 Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.1 Обладает фундаментальными знаниями в области прикладного программирования и информационных технологий	Знает программные средства и информационные технологии, используемые в высокопроизводительных вычислениях

Код и название компетенции	Код и название индикатора компетенции	Знания, умения, навыки
ОПК-3 Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.2 Умеет использовать их в профессиональной деятельности	Умеет использовать полученные знания для решения задач профессиональной деятельности
ОПК-3 Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.3 Имеет практический опыт применения программных средств, используемых при построении математических моделей в естественных науках	Владеет навыками применения и разработки программных средств в сфере высокопроизводительных вычислений

12. Объем дисциплины в зачетных единицах/час:

3/108

Форма промежуточной аттестации:

Экзамен

13. Трудоемкость по видам учебной работы

Вид учебной работы		Трудоемкость	
		Всего	По семестрам
			2 семестр
Аудиторные занятия		48	48
в том числе:	лекции	16	16
	практические		
	лабораторные	32	32
Самостоятельная работа		24	24
в том числе: курсовая работа (проект)			
Форма промежуточной аттестации (экзамен – 36 час.)		36	36
Итого:		108	108

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
1	Введение в высокопроизводительные вычисления	Роль и значение высокопроизводительных вычислений в современном мире. Производительность вычислительных систем. Закон Амдала	https://edu.vsu.ru/course/view.php?id=3968
2	Технологии параллельного программирования	Параллельное программирование с использованием MPI. Структура MPI-программы. Сообщения, их передача и прием. Синхронное и асинхронное взаимодействие. Коллективный обмен данными. Виды коллективного обмена, барьеры, широковещательная рассылка данных. Коммуникаторы и топологии. Производные типы данных. Компиляция и отладка MPI-программ. Система программирования OpenMP. Распределенные вычисления с использованием GRID	https://edu.vsu.ru/course/view.php?id=3968
3	Параллельные алгоритмы	Информационный граф алгоритма. Показатели эффективности параллельного алгоритма. Умножение матрицы на вектор. Матричное умножение	https://edu.vsu.ru/course/view.php?id=3968

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
4	Ведение в анализ больших данных	<p>Основные определения, термины, задачи анализа больших данных.</p> <p>Понятие «Data Mining».</p> <p>Когнитивный анализ данных. Обзор источников информации для Big Data. Открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований; доступ к закрытой информации)</p>	https://edu.vsu.ru/course/view.php?id=3968
5	Технологии хранения и обработки больших данных	<p>Обзор технологий хранения больших данных. Базы данных. Системы управления базами данных. Модели данных. NoSQL СУБД. Технология Map-Reduce. Ее реализация Hadoop. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных</p>	https://edu.vsu.ru/course/view.php?id=3968

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
6	Статистические методы анализа данных	Содержание темы. Основные понятия теории вероятностей и математической статистики. Методы анализа данных: дескриптивная статистика, параметрические, непараметрические, номинальные методы (корреляционный, регрессионный, дисперсионный анализы, кластерный, дискриминантный, факторный анализы). Обзор современных программных средства анализа данных: Statistica, SPSS, Excel, R-Studio и другие; их преимущества и недостатки. Языки Python и R. Синтаксис языка R, основные типы данных	https://edu.vsu.ru/course/view.php?id=3968

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
1	Введение в высокопроизводительные вычисления	2		0	6	8
2	Технологии параллельного программирования	2		6	8	16
3	Параллельные алгоритмы	2		6	12	20
4	Введение в анализ больших данных	2		6	12	20
5	Технологии хранения и обработки больших данных	4		8	10	22

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
6	Статистические методы анализа данных	4		8	10	22
		16	0	34	58	108

14. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины складывается из аудиторной работы (учебной деятельности, выполняемой под руководством преподавателя) и внеаудиторной работы (учебной деятельности, реализуемой обучающимся самостоятельно).

Аудиторная работа состоит из работы на лекциях и выполнения практических (или лабораторных) заданий в объёме, предусмотренном учебным планом. Лекция представляет собой последовательное и систематическое изложение учебного материала, направленное на знакомство обучающихся с основными понятиями и теоретическими положениями изучаемой дисциплины. Лекционные занятия формируют базу для практических (или лабораторных) занятий, на которых полученные теоретические знания применяются для решения конкретных практических задач. Обучающимся для успешного освоения дисциплины рекомендуется вести конспект лекций и практических (лабораторных) занятий.

Самостоятельная работа предполагает углублённое изучение отдельных разделов дисциплины с использованием литературы, рекомендованной преподавателем, а также конспектов лекций, презентационным материалом (при наличии) и конспектов практических (лабораторных) занятий. В качестве плана для самостоятельной работы может быть использован раздел 13.1 настоящей рабочей программы, в котором зафиксированы разделы дисциплины и их содержание. В разделе 13.2 рабочей программы определяется количество часов, отводимое на самостоятельную работу по каждому разделу дисциплины. Больше количество часов на самостоятельную работу отводится на наиболее трудные разделы дисциплины. Для самостоятельного изучения отдельных разделов дисциплины используется перечень литературы и других ресурсов, перечисленных в пунктах 15 и 16 настоящей рабочей программы.

Успешность освоения дисциплины определяется систематичностью и глубиной аудиторной и внеаудиторной работы обучающегося.

При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей, вовремя подключаться к online занятиям, ответственно подходить к заданиям для самостоятельной работы.

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

№ п/п	Источник
1	Бутаков, Н. А. Обработка больших данных с Apache Spark : учебно-методическое пособие / Н.А. Бутаков, М.В. Петров, Д. Насонов ; Министерство образования и науки Российской Федерации ; Университет ИТМО .— Санкт-Петербург : Университет ИТМО, 2019 .— 52 с. : ил. — Библиогр. в кн .— http://biblioclub.ru/ .— <URL: http://biblioclub.ru/index.php?page=book&id=566771 >.

№ п/п	Источник
2	Уэс, М. Python и анализ данных [Электронный ресурс] / Уэс М. ; Пер. с англ. Слинкин А.А. — Москва : ДМК Пресс, 2015 .— 482 с. — Книга из коллекции ДМК Пресс - Информатика .— ISBN 978-5-97060-315-4 .— <URL: http://e.lanbook.com/books/element.php?pl1_id=73074 >.

б) дополнительная литература:

№ п/п	Источник
1	Юре, Л. Анализ больших наборов данных [Электронный ресурс] / Юре Л. , Ананд Р. , Джеффри Д. У. ; Пер. с англ. Слинкин А.А. — Москва : ДМК Пресс, 2016 .— 498 с. — Книга из коллекции ДМК Пресс - Информатика .— ISBN 978-5-97060-190-7 .— <URL: https://e.lanbook.com/book/93571 >.

в) информационные электронно-образовательные ресурсы:

№ п/п	Источник
1	Электронный университет ВГУ https://edu.vsu.ru
2	ЭБС «Лань» https://e.lanbook.com/
3	«Университетская библиотека online» https://biblioclub.ru/
4	«Консультант студента» http://www.studmedlib.ru/
5	Электронный университет ВГУ: https://edu.vsu.ru/

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	Юре, Л. Анализ больших наборов данных [Электронный ресурс] / Юре Л. , Ананд Р. , Джеффри Д. У. ; Пер. с англ. Слинкин А.А. — Москва : ДМК Пресс, 2016 .— 498 с. — Книга из коллекции ДМК Пресс - Информатика .— ISBN 978-5-97060-190-7 .— <URL: https://e.lanbook.com/book/93571 >.

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):

При реализации дисциплины могут использоваться технологии электронного обучения и дистанционные образовательные технологии на базе портала edu.vsu.ru, а также другие доступные ресурсы сети Интернет.

18. Материально-техническое обеспечение дисциплины:

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 477

Учебная аудитория: специализированная мебель, ноутбук HP Pavilion Dv9000-er, мультимедийный проектор, экран

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 479

Учебная аудитория: специализированная мебель, компьютер преподавателя i5-8400-2,8ГГц, монитор с ЖК 19», мультимедийный проектор, экран

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 505п

Учебная аудитория: специализированная мебель, компьютер преподавателя i5-3220-3.3ГГц, монитор с ЖК 17», мультимедийный проектор, экран

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 292

Учебная аудитория: специализированная мебель, компьютер преподавателя Pentium-G3420-3,2ГГц, монитор с ЖК 17», мультимедийный проектор, экран. Система для видеоконференций Logitech ConferenceCam

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 297

Учебная аудитория: специализированная мебель, компьютер преподавателя i3-3240-3,4ГГц, монитор с ЖК 17», мультимедийный проектор, экран

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 380

Учебная аудитория: специализированная мебель, компьютер преподавателя i3-3240-3,4ГГц, монитор с ЖК 17», мультимедийный проектор, экран

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 290

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i7-7800x-4ГГц, мониторы ЖК 27» (12 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 291

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-3220-3,3ГГц, мониторы ЖК 19» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 293

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-8100-3,6ГГц, мониторы ЖК 22» (17 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 295

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-9100-3,6ГГц, мониторы ЖК 24» (14 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 382

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i5-9600KF-3,7ГГц, мониторы ЖК 24» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 383

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i7-9700F-3ГГц, мониторы ЖК 27» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 384

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-2120-3,3ГГц, мониторы ЖК 22» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 385

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-2120-3,3ГГц, мониторы ЖК 19» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 301п

Компьютерный класс: специализированная мебель, персональные компьютеры на базе

i3-2120-3,3ГГц, мониторы ЖК 17» (15 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 303п

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-8100-3,9ГГц, мониторы ЖК 24» (13 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 314п

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-7100-3,6ГГц, мониторы ЖК 19» (16 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

394018, г. Воронеж, площадь Университетская, д. 1, ауд. 316п

Компьютерный класс: специализированная мебель, персональные компьютеры на базе i3-9100-3,6ГГц, мониторы ЖК 19» (30 шт.), мультимедийный проектор, экран.

ПО: ОС Windows v.7, 8, 10, Набор утилит (архиваторы, файл-менеджеры), LibreOffice v.5-7, Пакет MPICH реализации параллельных вычислений, Foxit PDF Reader

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Разделы дисциплины (модули)	Код компетенции	Код индикатора	Оценочные средства для текущей аттестации
1	Разделы 1-6	ОПК-2	ОПК-2.1	Лабораторные работы
2	Разделы 1-6	ОПК-2	ОПК-2.2	Лабораторные работы
3	Разделы 1-6	ОПК-2	ОПК-2.3	Лабораторные работы
4	Разделы 1-6	ОПК-3	ОПК-3.1	Лабораторные работы
5	Разделы 1-6	ОПК-3	ОПК-3.2	Лабораторные работы
6	Разделы 1-6	ОПК-3	ОПК-3.3	Лабораторные работы

Промежуточная аттестация

Форма контроля - Экзамен

Оценочные средства для промежуточной аттестации

Перечень вопросов к экзамену

20 Типовые оценочные средства и методические материалы, определяющие процедуры

оценивания

20.1 Текущий контроль успеваемости

Текущий контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств:

- лабораторная работа.

Перечень лабораторных работ.

1. Параллельное программирование с использованием MPI.
2. Система программирования OpenMP.
3. Распределенные вычисления с использованием GRID технологий.
4. Матричное умножение.
5. Сортировка.
6. Базы данных. Системы управления базами данных.
7. NoSQL СУБД.
8. Технология Map-Reduce. Ее реализация Hadoop.
9. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных.
10. Методы анализа данных: дескриптивная статистика, параметрические, непараметрические, номинальные методы (корреляционный, регрессионный, дисперсионный анализы, кластерный, дискриминантный, факторный анализы).
11. Обзор современных программных средства анализа данных: Statistica, SPSS, Excel, R-Studio.
12. Языки Python и R. Синтаксис языка R, основные типы данных.

20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств:

- перечень вопросов к экзамену.

Перечень вопросов к экзамену

1. Роль и значение высокопроизводительных вычислений в современном мире. Производительность вычислительных систем. Закон Амдала.
2. Параллельное программирование с использованием MPI. Структура MPI-программы.
3. Сообщения, их передача и прием. Синхронное и асинхронное взаимодействие.
4. Коллективный обмен данными. Виды коллективного обмена, барьеры, широковещательная рассылка данных.
5. Коммуникаторы и топологии.
6. Производные типы данных.
7. Компиляция и отладка MPI-программ.
8. Система программирования OpenMP.
9. Распределенные вычисления с использованием GRID технологий.
10. Информационный граф алгоритма.
11. Показатели эффективности параллельного алгоритма.
12. Умножение матрицы на вектор.
13. Матричное умножение.
14. Сортировка.

Приведённые ниже задания рекомендуется использовать при проведении диагностических работ для оценки остаточных знаний по дисциплине

тестовые задания - 1 балл

1. Компиляция программы в рамках стандарта MPI может быть выполнена с помощью команды:

- a) mpirun
- b) mpifort**
- c) mpicxx
- d) mpicc

2. Получить значение размера коммуникатора можно с помощью подпрограммы MPI:

- a) MPI_Comm_size**
- b) MPI_Comm_rank
- c) MPI_Init
- d) MPI_Finalize

3. В рамках операции редукции в стандарте MPI можно выполнить

- a) суммирование элементов массива**
- b) умножение двух матриц
- c) **определение максимального элемента**
- d) **перемножение элементов массива**

4. Операция широковещательной рассылки

- a) пересылает одни и те же данные всем процессам**
- b) выполняет частичную редукцию
- c) пересылает всем остальным процессам разные фрагменты данных
- d)

5. Подпрограмма MPI_Gather:

- a) выполняет широковещательную рассылку
- b) выполняет сбор данных**
- c) **является операцией коллективного обмена данными**
- d)

6. Подпрограмма MPI_Scatter:

- a) выполняет распределение данных между процессами
- b) выполняет широковещательную рассылку
- c) является операцией синхронизации
- d)

7. Какие из следующих операций выполняют сбор данных на одном процессе

- a) MPI_Gather**
- b) MPI_Gatherv**
- c) MPI_Allgather
- d)

8. Какие их перечисленных операций относятся к классу векторных

- a) MPI_Alltoall
- b) MPI_Alltoallv**
- c) MPI_Allgather
- d) **MPI_Allgatherv**

9. Подпрограмма MPI_Alltoall:

- a) выполняет операцию частичной редукции
- b) выполняет операцию полной редукции
- c) выполняет операцию синхронизации

d) выполняет передачу данных по схеме «каждый-всем»

10. Коммуникатором в MPI называется:

- a) **группа процессов, наделенная общим контекстом обмена данными**
- b) подпрограмма, выполняющая передачу или прием сообщения
- c) подпрограмма для обработки сигналов, поступающих от операционной системы
- d)

11. Для создания коммуникатора предусмотрена команда

- a) MPI_Comm_size
- b) **MPI_Comm_create**
- c) MPI_Comm_rank
- d)

12. Создать группу коммуникаторов можно с помощью команды

- a) **MPI_Group_excl**
- b) **MPI_Group_incl**
- c) **MPI_Group_difference**
- d) MPI_Group_free

13. Подпрограмма MPI_Comm_size:

- a) является операцией двухточечного обмена
- b) **возвращает количество процессов в указанном коммуникаторе**
- c) задает количество процессов в указанном коммуникаторе
- d) проверяет возможность создания коммуникатора с указанным размером

14. Определить количество процессов в коммуникаторе можно с помощью:

- a) **MPI_Comm_size**
- b) MPI_Comm_rank
- c) MPI_Comm_create
- d)

15. В MPI существуют следующие типы двухточечных обменов:

- a) **стандартный обмен**
- b) **обмен по готовности**
- c) **синхронный обмен**
- d) **обмен с буферизацией**

16. Стандартная блокирующая двухточечная передача выполняется

- a) **MPI_Send**
- b) MPI_Ssend
- c) MPI_Bsend
- d) MPI_Rsend

17. Двухточечная передача с буферизацией выполняется функцией

- a) MPI_Send
- b) MPI_Ssend
- c) **MPI_Bsend**
- d) MPI_Rsend

18. Неблокирующий вариант операций передачи сообщений существует для:

- a) **стандартной операции обмена данными**
- b) **обмена с буферизацией**
- c) **обмена «по готовности»**
- d)

19. Неблокирующий прием в MPI выполняется :

- a) **MPI_Irecv**
- b) MPI_Recv
- c) MPI_Reduce_Scatter
- d)

20. Подпрограмма MPI_Wait предназначена для:

- a) **блокирующей проверки выполнения обмена**
- b) неблокирующей проверки выполнения обмена
- c) приостановки выполнения программы на заданный период времени
- d)

21. Подпрограмма MPI_Test предназначена для:

- a) блокирующей проверки выполнения обмена
- b) **неблокирующей проверки выполнения обмена**
- c) приостановки выполнения программы на заданный период времени
- d)

22. Подпрограмма MPI_Testall выполняет проверку:

- a) завершения любой операции обмена из некоторого множества
- b) **завершения всех обменов**
- c) завершения одного обмена
- d)

23. Подпрограмма MPI_Testany выполняет проверку:

- a) **завершения любой операции обмена из некоторого множества**
- b) завершения всех обменов
- c) завершения одного обмена
- d)

24. Неблокирующая передача с буферизацией выполняется :

- a) **MPI_IbSEND**
- b) MPI_BSEND
- c) MPI_Test
- d) MPI_Testall

25. Средством разработки многопоточных программ является:

- a) **OpenMP**
- b) PVM
- c) MPI
- d)

26. Какие виды зависимости операторов в программе существуют

- a) **зависимость по выходу**
- b) **потокосная зависимость**
- c) **антизависимость**
- d)

27. Какие ключевые слова OpenMP можно использовать после директивы parallel

- a) **firstprivate**
- b) **lastprivate**
- c) **reduction**
- d) runtime

28. Какие ключевые слова OpenMP можно использовать после директивы single

- a) **private**
- b) **firstprivate**

- c) **nowait**
- d) **copyprivate**

29. Какой термин применяется для обозначения структурированных и неструктурированных данных значительного объема и многообразия, эффективно обрабатываемых горизонтально масштабируемыми программными инструментами?

- a) **большие данные**
- b) суперкомпьютерные вычисления
- c) интеллектуальный анализ данных
- d) облачные технологии

30. Какие технологии не относятся к технологиям анализа больших данных

- a) **конвейерные вычисления**
- b) системы управления базами данных категории NoSQL
- c) алгоритмы MapReduce
- d) методы интеллектуального анализа данных

31. Какой термин не входит в набор признаков VVV, изначально выбранных для характеристики больших данных?

- a) **континуальность**
- b) объем
- c) скорость
- d) разнообразие

2) задания с коротким ответом — 2 балла

1. Назовите метод исследования больших данных, при котором изучаемая система заменяется моделью, с достаточной точностью описывающей реальную систему, с которой проводятся эксперименты с целью получения информации об этой системе.

Ответ: имитационное моделирование

2. Какой аббревиатурой принято обозначать модель вычислительной системы – параллельную машину с произвольным доступом к памяти?

Ответ: PRAM

3. Как называется величина, равная отношению $T_1(N)/T_p(N)$, где p – количество вычислительных узлов, $T_1(N)$ и $T_p(N)$ – время работы вычислительного алгоритма на системе с одним и p вычислительными узлами соответственно?

Ответ: ускорение **или** ускорение алгоритма

4. Как называется система управления базами данных, существенно отличающихся от традиционных реляционных СУБД с доступом к данным средствами языка SQL за счет решения проблемы масштабируемости и доступности путем полного или частичного отказа от требований атомарности и согласованности данных.

Ответ: NoSQL

5. Какую аббревиатуру используют для обозначения отдельного устройства вычислительной системы, способное выполнять графический рендеринг, а также высокопроизводительные вычисления наряду с центральным процессором?

Ответ: GPU

3) задания с развернутым ответом

1. Время работы последовательной версии некоторого алгоритма B равно $T_1(N) = 2^N \tau$, где N – размер входных данных, τ – время выполнения одной вычислительной операции. В предположении, что алгоритм допускает максимальное распараллеливание, т.е. время работы на вычислительной системе с p процессорами равно $T_p(N) = T_1(N) / p$, вычислите время работы алгоритма B в следующем случае: $N = 32$, $p = 4$.

2. Время работы последовательной версии некоторого алгоритма B равно $T_1(N) = 2^N \tau$, где N – размер входных данных, τ – время выполнения одной вычислительной операции. В предположении, что алгоритм допускает максимальное распараллеливание, т.е. время работы на вычислительной системе с p процессорами равно $T_p(N) = T_1(N) / p$, вычислите время работы алгоритма B в следующем случае: $N = 32$, $p = 16$.

3. Пусть доля последовательных вычислений f в некоторой программе равна $1/10$. Вычислите максимальное ускорение программы $(S_p)_{\max}$ на вычислительной системе с p процессорами с учетом закона Амдала.

4. Пусть доля последовательных вычислений f в некоторой программе равна $1/100$. Вычислите максимальное ускорение программы S_∞ с учетом закона Амдала.

5. Покажите, что условия Бернштейна не являются необходимыми условиями коммутативности двух операторов

6. На экзамене студент утверждает, что в рамках коммутатора любой процесс может отправить с помощью функции `MPI_Send()` сообщение любому процессу, в том числе и самому себе. Прав ли студент?

7. Пусть организуется попарный обмен сообщениями в рамках коммуникатора MPI_COMM_WORLD, и количество процессов в параллельной программе равно N. Запишите множество рангов, доступных для использования в качестве рангов процесса-получателя в MPI_Send();

8. Пусть организуется попарный обмен сообщениями в рамках коммуникатора MPI_COMM_WORLD, и количество процессов в параллельной программе равно N. Запишите множество рангов, доступных для использования в качестве рангов процесса-отправителя в MPI_Recv().

Ответы

1. $T_p(N) = 2^{(30)} \tau$

2. $T_p(N) = 2^{(28)} \tau$

3. $10p/(p+9)$

4. $S_{\infty} = 100$

5. Рассмотрим следующую пару операторов (o1, o2): o1 : a = a + 1; o2 : a = a - 1;

Ясно, что все три условия Бернштейна в данном случае нарушены. Тем не менее, результат работы программ, содержащих последовательности (o1, o2) и (o2, o1), будет одинаков для всех возможных значений переменной a. Таким образом, условия Бернштейна не являются необходимыми условиями коммутативности двух операторов.

6. Да, в среде MPI процесс коммуникатора может отправить сообщение любому процессу, принадлежащему тому же коммуникатору, в том числе и себе. Тем не менее, указанной возможности в программе желательно избегать, т.к. в данном случае возможно появление ситуации «тупика».

7. $rank \in \{0,1,\dots,N - 1\} \cup \{MPI_PROC_NULL\}$

8. $rank \in \{0,1,\dots,N - 1\} \cup \{MPI_PROC_NULL\} \cup \{MPI_ANY_SOURCE\}$

Критерии и шкалы оценивания при проведении промежуточной аттестации

Шкала оценивания (суммарный % правильных ответов)	Критерии оценивания
«отлично» (85% –100 %)	Оценка «отлично» выставляется, если студент правильно ответил на теоретические вопросы. Показал отличные знания в рамках учебного материала. Правильно выполнил практическое задание. Показал отличные умения и владения навыками применения полученных знаний и умений при решении задач в рамках учебного материала. Ответил на все дополнительные вопросы.
«хорошо» (71% – 84 %)	Оценка «хорошо» выставляется, если студент с небольшими неточностями ответил на теоретические вопросы. Показал хорошие знания в рамках учебного материала. С небольшими неточностями выполнил практические задания. Показал хорошие умения и владения навыками применения полученных знаний и умений при решении задач в рамках учебного материала. Ответил на большинство дополнительных вопросов.
«удовлетворительно»	Оценка «удовлетворительно» выставляется, если

<p>(51% – 70 %)</p>	<p>студент с существенными неточностями ответил на теоретические вопросы. Показал удовлетворительные знания в рамках учебного материала. С существенными неточностями выполнил практические задания. Показал удовлетворительные умения и владения навыками применения полученных знаний и умений при решении задач в рамках учебного материала. Допустил много неточностей при ответе на дополнительные вопросы.</p>
<p>«неудовлетворительно» (0% – 50 %)</p>	<p>Оценка «неудовлетворительно» выставляется, если студент при ответе на теоретические вопросы и при выполнении практических заданий продемонстрировал недостаточный уровень знаний и умений при решении задач в рамках учебного материала. При ответах на дополнительные вопросы было допущено множество неправильных ответов.</p>