

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

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ:

Директор

А. В. Замятин

« 19 » мая 20 22 г.

Рабочая программа дисциплины

Технологии высокопроизводительной обработки больших данных

по направлению подготовки / специальности

10.05.01 Компьютерная безопасность

Направленность (профиль) подготовки / специализация:

Анализ безопасности компьютерных систем

Форма обучения

Очная

Квалификация

Специалист по защите информации

Год приема

2022

Код дисциплины в учебном плане: Б1.В.ДВ.01.01.04

СОГЛАСОВАНО:

Руководитель ОП

В.Н. Треньякаев

Председатель УМК

С.П. Сущенко

Томск – 2022

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

– ОПК-1 – Способен оценивать роль информации, информационных технологий и информационной безопасности в современном обществе, их значение для обеспечения объективных потребностей личности, общества и государства.

– ОПК-2 – Способен применять программные средства системного и прикладного назначений, в том числе отечественного производства, для решения задач профессиональной деятельности.

– ОПК-3 – Способен на основании совокупности математических методов разрабатывать, обосновывать и реализовывать процедуры решения задач профессиональной деятельности.

– ОПК-14 – Способен проектировать базы данных, администрировать системы управления базами данных в соответствии с требованиями по защите информации.

– ПК-2 – Способен разрабатывать требования к программно-аппаратным средствам защиты информации компьютерных систем и сетей.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-1.1 Учитывает современные тенденции развития информационных технологий в своей профессиональной деятельности.

ИОПК-2.3 Формулирует предложения по применению программных средств системного и прикладного назначений, в том числе отечественного производства, используемых для решения задач профессиональной деятельности.

ИОПК-3.2 Осуществляет применение основных понятий, фактов, концепций, принципов математики и информатики для решения задач профессиональной деятельности.

ИОПК-14.1 Понимает модели и структуры данных, физические модели баз данных, принципы организации и методы проектирования баз данных, языки и системы программирования баз данных.

ИПК-2.2 Разрабатывает математические модели, реализуемые в средствах защиты информации.

ИПК-2.3 Проводит исследования с целью нахождения наиболее целесообразных практических решений по обеспечению защиты информации.

2. Задачи освоения дисциплины

– обучить студентов различным методам организации высокопроизводительных вычислений;

– сформировать у студентов навыки использования различных программных инструментов для организации параллельных вычислений.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль по выбору "Введение в искусственный интеллект" модуля «Профессиональные модули по выбору».

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. Введение

Актуальность, базовая терминология и тенденции развития. Параллелизм компьютерных вычислений.

Тема 2. Архитектура вычислительных систем

Архитектура высокопроизводительных вычислительных систем. Классификация вычислительных систем.

Тема 3. Облачные технологии.

Облачные технологии, их свойства и типы

Тема 4. MapReduce

Технология вычислений MapReduce

Тема 5. Распределённые файловые системы

Тема 6. Программирование для высокопроизводительных вычислений.

Методы программирования для высокопроизводительных вычислений. Методология проектирования параллельных алгоритмов

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, вопросов по лекционному материалу, контроля выполнения практических работ, и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Экзамен по окончании изучения курса проводится в письменной или устной форме.

Примерный перечень теоретических вопросов:

1. Виды параллелизма компьютерных вычислений.
2. Классификацию вычислительных систем Флинна.
3. Классификация MIMD систем.
4. Свойства и типы облачных технологий.
5. Технология распределённых вычислений MapReduce.
6. Реализации MapReduce: сходства и различия.
7. Распределённые файловые системы.
8. Распределённая файловая система Google File System.
9. Распределённая файловая система Hadoop distributed file system.
10. Методология организации параллельных вычислений для SIMD, MIMD архитектур.
11. Показатели качества параллельных методов.
12. OpenMP и MPI: назначение каждой из технологий и их сравнительный анализ.

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.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle»

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине (Приложение 1).

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. – 400 с.

– A Ohri. R for Cloud Computing electronic resource : An Approach for Data Scientists. – New York, NY : Springer New York : Imprint: Springer, 2014. – 336 с.

– Zaigham Mahmood. Cloud Computing: Methods and Practical Approaches, – London : Springer London : Imprint: Springer, 2013. – 259 с.

– К. Ю. Богачев. Основы параллельного программирования / Москва БИНОМ. Лаборатория знаний, 2013. – 389 с.

– В. П. Гергель. Теория и практика параллельных вычислений: учебное пособие / Москва : Интернет-Университет Информационных Технологий, 2016. – 426 с.

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. - 280 с.

– Xiaolin Li, Judy Qiu. Cloud Computing for Data-Intensive Applications electronic resource – New York, NY : Springer New York : Imprint: Springer, 2014. – 226 с.

– Zaigham Mahmood. Cloud Computing electronic resource: Challenges, Limitations and R&D Solutions – New York, NY : Springer New York : Imprint: Springer, 2014. – 321 с.

– А. В. Линева, Д. К. Боголепов, С. И. Бастраков. Технологии параллельного программирования для процессоров новых архитектур – Нижегородский гос. ун-т им. Н. И. Лобачевского – Москва: Изд-во Московского университет, 2010. – 572 с.

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://swwsweb.ru/cloud-computing-basic-concepts-problems.html>

– MapReduce Tutorial: 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

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Visual Studio Community Edition 2013 (C++), библиотека классов, реализующая технологию MapReduce (например, <https://github.com/cdmh/mapreduce>). Все используемые программные продукты являются свободно распространяемыми.

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 занятий семинарского типа, 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.