

**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

	Рабочая программа дисциплины (модуля)
по дисциплине:	Проектирование высоконагруженных систем
по направлению:	Прикладная математика и информатика
профиль подготовки:	А1360: Передовые методы искусственного интеллекта Физтех-школа Прикладной Математики и Информатики кафедра технологий цифровой трансформации
курс:	4
квалификация:	бакалавр

Семестр, формы промежуточной аттестации: 8 (весенний) - Дифференцированный зачет

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Всего часов: 135, всего зач. ед.: 3

Программу составил: О.В. Бунин, ассистент

Программа обсуждена на заседании кафедры технологий цифровой трансформации 07.02.2023

Аннотация

В курсе изучается, что можно считать высоконагруженной системой и в каких единицах измеряется нагрузка. Описываются различные архитектуры веб-серверов, устройство типичного веб-сайта. Рассказывается о методах подключения динамического содержимого.

1. Цели и задачи

Цель дисциплины

- изучение инструментов и технологий программирования для создания высоконагруженных сервисов и веб-приложений.

Задачи дисциплины

- совершенствование знаний по веб-разработке;
- получение опыта разработки высоконагруженных приложений;
- получение опыта практической работы с большими базами данных;
- получение опыта проектирования больших систем;
- формирование культуры комплексного подхода к выбору и построению архитектуры.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-3 Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде	УК-3.1 Способен устанавливать разные виды коммуникации (учебную, научную, деловую, неформальную и др.)
	УК-3.2 Взаимодействует с другими членами команды для достижения поставленной задачи
ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач	ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности
	ОПК-4.2 Знает основные источники научно-технической и (или) технологической информации в области профессиональной деятельности
	ОПК-4.3 Умеет составлять аннотации, рефераты, библиографические перечни и обзоры информации в области своей профессиональной деятельности
	ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации
ОПК-5 Способен участвовать в проведении фундаментальных и прикладных исследований и разработок, самостоятельно осваивать новые теоретические, в том числе, математические методы исследований и работать на современной экспериментальной научно-исследовательской, измерительно-аналитической и технологической аппаратуре	ОПК-5.1 Способен решать поставленные задачи в области теоретических и экспериментальных исследований и разработок
	ОПК-5.2 Обладает способностью к освоению новых знаний на основе изучения литературы, научных статей и других источников
	ОПК-5.3 Способен к профессиональной эксплуатации современной экспериментальной научно-исследовательской (измерительно-аналитической и технологической) аппаратуры
ПК-1 Способен ставить, формализовывать и	ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности

решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели
	ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации
	ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива
	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:

- C10k problem;
- устройства популярных веб-серверов;
- реляционные БД и различия в устройстве движков СУБД;
- принципы устройства нереляционных СУБД;
- виды репликаций;
- механизмы кластеризации БД;
- механизмы кеширования;
- асинхронные фреймворки;
- механизмы отдачи статики и организации CDN;
- очереди сообщений;
- организацию и инструменты для полнотекстового поиска;
- принцип работы шаблонизаторов и популярные библиотеки;
- устройство файловых систем.

уметь:

- настраивать веб-сервера;
- обеспечивать безопасность хранения и передачи данных между серверами;
- проектировать шардирование данных;
- настраивать репликацию и кластеризацию БД;
- проектировать архитектуру высоконагруженных приложений и сервисов;
- использовать системы кеширования;
- использовать асинхронные фреймворки для разработки высоконагруженных приложений;
- создавать CDN;
- организовывать полнотекстовый поиск;
- обеспечивать балансировку нагрузки;
- настраивать раздачу статики.

владеть:

- скриптовыми языками командных оболочек;
- инструментами администрирования и конфигурирования БД;
- языками программирования Python, PHP, Javascript.

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

		Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.
--	--	---

№	Тема (раздел) дисциплины	Лекции	Семинары	Лаборат. работы	Самост. работа
1	Трёхзвенная архитектура	3	3		8
2	Кеширование	3	3		7
3	Использование толстого клиента	3	3		8
4	Деграция функциональности	3	3		7
5	Вертикальное и горизонтальное масштабирование	3	3		8
6	Масштабирование во времени	3	3		7
7	Масштабирование баз данных	3	3		8
8	Специализированные сервера	3	3		7
9	Антипаттерны	3	3		8
10	Конвейер	3	3		7
Итого часов		30	30		75
Подготовка к экзамену		0 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

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

Семестр: 8 (Весенний)

1. Трёхзвенная архитектура

Фронтенд, бекенд, система хранения. Отдача статического контента, буферизация запросов, масштабирование бекендов, обслуживание медленных клиентов.

2. Кеширование

Кеширование в браузере, HTML-блоков, данных, HTML-страниц. Единый кеш для всех бекендов, проблема инвалидации кеша, проблема холодного старта, целесообразность применения кеша. Race condition, проблема одновременного перестроения кешей. Тегирование кеша.

3. Использование толстого клиента

Single Page Application, проблема антишквала, умная балансировка.

4. Деграция функциональности

Снижение уровня системной структурированности. Функционал веб-сайта можно последовательно наращивать для разных групп пользователей.

5. Вертикальное и горизонтальное масштабирование

Максимальная независимость компонент. Гомогенные одноранговые сервера. Отсутствие единой точки отказа. Закон Амдала. Балансировка фронтендов, балансировка бекендов.

6. Масштабирование во времени

Отложенные вычисления, стадии обработки запроса. Демоны. Асинхронная обработка. Очереди, FIFO, LIFO. Интеркоммуникация сервисов.

7. Масштабирование баз данных

Требования ACID, MVCC. CAP-теорема, согласованность, доступность, устойчивость к разделению. Репликация, вертикальный шардинг, функциональное разделение баз данных. Горизонтальный шардинг, виртуальные шарды, центральный диспетчер, партиционирование. Денормализация, первая, вторая и прочие нормальные формы.

8. Специализированные сервера

Круг задач, которые должны выполнять серверы, многообразен и сложен. Чтобы приспособиться к возрастающим потребностям пользователей, серверы в больших сетях стали специализированными (specialized). Например, в сети Windows NT существуют различные типы серверов.

9. Антипаттерны

распространённый подход к решению класса часто встречающихся проблем, являющийся неэффективным, рискованным или непродуктивным. В отличие от шаблона проектирования.

10. Конвейер

Работы по проектированию конвейеров выполняются в несколько этапов:

Выводится и утверждается техническое задание.

Осуществляются конструкторские расчеты.

Разрабатывается проект.

Чертежи сдаются Клиенту.

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

учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

1. Моделирование и проектирование информационных систем с использованием методологии ARIS [Текст] : Лабораторный практикум : учеб. пособие для вузов / Е. Б. Степанова, В. А. Верещагин ; М-во образования и науки Рос. Федерации, Фед. агентство по образованию, Моск. физ.-техн. ин-т (гос. ун-т), Фак. инноваций и высоких технологий .— М. : Изд-во МФТИ, 2007 .— 118 с.

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)

Для контроля и коррекции знаний обучающиеся могут использовать компьютерное тестирование.

Желательно ознакомиться с курсом Бунина О.В. " Разработка высоконагруженных систем." По материалам конференции HighLoad++ 2010-2011, Год: 2012, издательство Олега Бунина.

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

Для успешного освоения данной дисциплины студенту необходимо:

- посещать лекции;
- выполнять задания, задаваемые преподавателем на лекциях;
- выполнить итоговое письменное задание по дисциплине, которое вносит основной вклад в изучение дисциплины, а также в итоговую оценку по данному курсу.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: AI360: Передовые методы искусственного интеллекта
Физтех-школа Прикладной Математики и Информатики
кафедра технологий цифровой трансформации
курс: 4
квалификация: бакалавр

Семестр, формы промежуточной аттестации: 8 (весенний) - Дифференцированный зачет

Разработчик: О.В. Бунин, ассистент

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-3 Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде	УК-3.1 Способен устанавливать разные виды коммуникации (учебную, научную, деловую, неформальную и др.)
	УК-3.2 Взаимодействует с другими членами команды для достижения поставленной задачи
ОПК-4 Способен осуществлять сбор и обработку научно-технической и (или) технологической информации для решения фундаментальных и прикладных задач	ОПК-4.1 Владеет методами научного поиска и интеллектуального анализа информации при решении задач профессиональной деятельности
	ОПК-4.2 Знает основные источники научно-технической и (или) технологической информации в области профессиональной деятельности
	ОПК-4.3 Умеет составлять аннотации, рефераты, библиографические перечни и обзоры информации в области своей профессиональной деятельности
	ОПК-4.4 Владеет навыками работы с компьютером и компьютерными сетями с целью получения, хранения и обработки научной (технической, технологической) информации
ОПК-5 Способен участвовать в проведении фундаментальных и прикладных исследований и разработок, самостоятельно осваивать новые теоретические, в том числе, математические методы исследований и работать на современной экспериментальной научно-исследовательской, измерительно-аналитической и технологической аппаратуре	ОПК-5.1 Способен решать поставленные задачи в области теоретических и экспериментальных исследований и разработок
	ОПК-5.2 Обладает способностью к освоению новых знаний на основе изучения литературы, научных статей и других источников
	ОПК-5.3 Способен к профессиональной эксплуатации современной экспериментальной научно-исследовательской (измерительно-аналитической и технологической) аппаратуры
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности
	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели
	ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации
	ПК-2.2 Способен планировать и проводить научные исследования самостоятельно или в качестве члена (руководителя) малого научного коллектива
	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 «Проектирование высоконагруженных систем» обучающийся должен:

знать:

- C10k problem;
- устройства популярных веб-серверов;
- реляционные БД и различия в устройстве движков СУБД;
- принципы устройства нереляционных СУБД;
- виды репликаций;
- механизмы кластеризации БД;
- механизмы кеширования;
- асинхронные фреймворки;
- механизмы отдачи статики и организации CDN;
- очереди сообщений;
- организацию и инструменты для полнотекстового поиска;
- принцип работы шаблонизаторов и популярные библиотеки;
- устройство файловых систем.

уметь:

- настраивать веб-сервера;
- обеспечивать безопасность хранения и передачи данных между серверами;
- проектировать шардирование данных;
- настраивать репликацию и кластеризацию БД;
- проектировать архитектуру высоконагруженных приложений и сервисов;
- использовать системы кеширования;
- использовать асинхронные фреймворки для разработки высоконагруженных приложений;
- создавать CDN;
- организовывать полнотекстовый поиск;
- обеспечивать балансировку нагрузки;
- настраивать раздачу статики.

владеть:

- скриптовыми языками командных оболочек;
- инструментами администрирования и конфигурирования БД;
- языками программирования Python, PHP, Javascript.

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

Перечень контрольных вопросов:

1. Трёхзвенная архитектура
2. Кеширование
3. Использование толстого клиента
4. Деграция функциональности
5. Вертикальное и горизонтальное масштабирование
6. Масштабирование во времени
7. Конвейер
8. Сервисно-ориентированная архитектура
9. Масштабирование баз данных
10. Специализированные сервера
11. Антипаттерны

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

1. Кто посещал мою страницу?

Как реализовать функцию "Кто посещал мою страницу за последние 24 часа" на крупном сайте знакомств (200 миллионов анкет, 500 миллионов просмотров анкет в сутки).

Как реализовать функцию невидимки (пользователь может купить невидимость на несколько минут)?

2. Хештеги

Спроектируйте хештеги в твиттере. Все сервисы, связанные с хештегом - поиск по хештегу, ленту хештега, рейтинг хештегов.

Уметь показать топ самых популярных хештегов среди друзей пользователя за день/неделю, которые он еще сегодня не смотрел.

Топ популярных хештегов показывается по регионам.

3. Товарные рекомендации

Система товарных рекомендаций, основанная на последних просмотренных, добавленных в корзину/вишлист и купленных товарах. Реальное время, один миллион просмотров в день, 10 тысяч товаров.

Не использовать машинное обучение.

4. Помощь страховой

Спроектировать систему, которая, получая данные с ГЛОНАС на автомобиле (скорее всего нужны ещё с десятков локальных датчиков в автомобиле), накапливает эти знания и после определённого количества "заходов в поворот боком" отправляет владельцу СМС с уведомлением о 20% повышении страховой премии.

Раз в двадцать минут сбрасывать информацию на сервер. Миллион машин и больше прямо сейчас ездят.

20 байт, одна минута, миллион машин, 60 минут -> 365 гигабайт в год.

5. Заказ такси

Спроектировать серверную часть заказа такси, при этом учитывать:

1. Реальное время доезда машины до заказчика, не просто километраж, а с учетом пробок;
2. Учитывать фильтры – курящий/не курящий, наличие детского сиденья и другие.

6. Eve Online

Спроектировать серверную часть онлайн игры, в которой может одновременно участвовать неограниченное множество игроков.

7. Mos.ru

Реализовать коннектор к большому количеству источников данных, все они разные:

- Некоторые из них готовы присылать данные, некоторые готовы только отвечать на запросы;
- Кто-то готов присылать изменения, кто-то готов присылать каждые данные только полностью слепком;
- Кто-то отвечает по 10 запросов в секунду, ограничивая количество запросов;
- Кто-то присылает данные в Excel или любой другой выгрузке.

9. Система видеонаблюдений за выборами

86000 участков, с каждого из них идёт три трансляции, которые автоматически записываются. Можно подключиться к любой трансляции и смотреть, можно прокрутить на любое время и смотреть.

Важно – надо быть готовым к ситуации, когда сто тысяч человек прибегают смотреть на один конкретный участок (а вдруг там Путин голосует?)

10. Онлайн-конференция

В конференции участвует очень много участников, надо организовать систему пуш-уведомлений о событиях, которые происходят на конференции. Пуш-уведомления появляются без перегрузки страницы.

В качестве усложнения - хорошо бы уведомления разбивать по интересам, чтобы каждый пользователь получал уведомления только по своим интересам.

Список дополнительных вопросов для экзамена

Буду задавать их тем, в чём ответе сомневаюсь.

1. Для чего нужна репликация? Когда она необходима? Когда она вредна?
2. Как измеряется эффективность кеширования?
3. Расскажите о тегировании кеша?
4. Какие три звена в трёхзвенной структуре?
5. За счёт чего увеличивается эффективность при использовании фронтеда? Приведите четыре аспекта?
6. Какие уровни кеширования вы знаете?
7. Что такое проблема непрогретого кеша? Как решается?
8. Какие способы инвалидировать кеш вы знаете?
9. Что за проблема - одновременное перестроение кешей? Как решается?
10. Как учитывать зависимости при кешировании и зачем?
11. Что такое проблема само-DDoS? Как с ней бороться?
12. Что такое толстый клиент?
13. Зачем используют комет-сервер?
14. Приведите три примера деградации функциональности?
15. Чем отличается горизонтальное от вертикального масштабирования?
16. Три бекенда подключены к толстым клиентам. Если бекенд перестаёт отвечать, то толстый клиент переключается на новый бекенд. В чём может быть проблема?
17. Какие способы масштабирования во времени вы знаете?
18. Как работает центральный диспетчер в шардировании? Как избежать проблемы неравномерного распределения по шардам?
19. Какие принципы горизонтального масштабирования вы знаете?
20. О чём говорит закон Амдала? Как он связан с горизонтальным масштабированием?
21. Как обеспечить отказоустойчивость звена фронтендов?
22. Как обеспечить отказоустойчивость бекендов?
23. Чем отличается SOA от монолита?
24. В чём плюсы монолитной разработки?
25. Как очереди используются в коммуникации между сервисами?
26. Что такое конвейер? Где используется?
27. Что такое горизонтальный шардинг?
28. Что такое виртуальные шарды?
29. Какие способы разбиения по шардам Вы знаете? В чём плюсы и минусы каждого?
30. Приведите примеры партиционирования базы данных.
31. Что такое денормализация и введение избыточности?
32. Чем плохо использование join'ов в SQL запросов в масштабированной системе?

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

Во время проведения дифференцированного зачета обучающиеся могут пользоваться программой дисциплины.