
УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол № 2

**Рабочая программа дисциплины (модуля)
«Системы управления базами данных»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Программная инженерия в
ритейле

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	8
6. Материально-техническое обеспечение	8
7. Методические и оценочные материалы	10

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Системы управления базами данных» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по направлению 02.04.01 Математика и компьютерные науки, профиль Программная инженерия в ретейле, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Системы управления базами данных» позволяет понять архитектурные особенности и компромиссы современных систем управления базами данных, научиться выбирать и эффективно использовать подходящие системы хранения данных под конкретную задачу, настраивать их для высокой производительности и отказоустойчивости, а также разрабатывать и эксплуатировать надёжные, масштабируемые и восстанавливаемые решения в распределённой среде.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Программная инженерия в ретейле и входит в обязательную часть Блока 1.

Дисциплина (модуль) доступна к изучению на 2 курсе в 3 семестре.

Цель изучения дисциплины (модуля): формирование у обучающихся глубокого понимания архитектуры современных распределённых систем управления базами данных, развитие навыков выбора, настройки и эксплуатации систем хранения данных с учётом требований к согласованности, доступности, производительности и отказоустойчивости.

Задачи изучения дисциплины (модуля):

- изучить принципы работы и компромиссы распределённых систем на основе теорем CAP и PACELC;
- освоить архитектуру и особенности ключевых NoSQL-систем: Redis, MongoDB, Cassandra, Kafka, ClickHouse, MinIO и векторных БД;
- научиться проектировать схемы, настраивать репликацию, шардирование, персистентность и индексы в различных системах управления базами данных;
- развить навыки развертывания, мониторинга и оптимизации производительности распределённых систем;
- овладеть практиками безопасной эволюции схем, управления жизненным циклом данных и построения отказоустойчивых пайплайнов.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- теоремы CAP и PACELC — компромиссы между согласованностью, доступностью и задержкой;
- классификацию NoSQL-систем: Key-Value, Document, Wide-Column, Graph, Vector, Object Storage, Message Log Storage;
- внутреннее устройство Redis: in-memory архитектура, персистентность (RDB, AOF), структуры данных (listpack, skiplist, hashtable, quicklist), транзакции, Lua-скрипты, кластеризация (хэш-слоты), Sentinel, алгоритм Redlock;
- устройство MongoDB: BSON, WiredTiger (MVCC, контрольные точки, сжатие), индексы, агрегационный pipeline, репликация (oplog, Raft), шардирование, change streams, эволюция схем;
- архитектуру Cassandra: LSM-деревья, Bloom-фильтры, модель данных (partition key, clustering key), уровни консистентности (ONE, QUORUM, ALL), gossip, hinted handoff, read repair, compaction, tombstones;
- принципы колоночного хранения в ClickHouse: векторизация, кодеки (Delta, Gorilla), MergeTree (ORDER BY, PRIMARY KEY, партиционирование),

- материализованные представления, шардирование;
- устройство объектных хранилищ (S3/MinIO): bucket, object, метаданные, модели консистентности, presigned URL, мультипартная загрузка, версионирование, lifecycle policies, erasure coding;
- log-структурированную архитектуру Kafka: партиции, сегменты, оффсеты, репликация (ISR), продюсеры (acks, идемпотентность), консьюмеры (consumer groups, ручной коммит), exactly-once (транзакции), Kafka Streams;
- основы векторных БД: эмбединги, метрики близости, индексы HNSW и IVF, сравнение pgvector и Milvus.

уметь:

- выбирать подходящую СУБД под задачу, обосновывая выбор CAP/PACELC;
- настраивать персистентность Redis (RDB, AOF) и оценивать влияние на производительность и время восстановления;
- реализовывать rate limiter, очередь задач и распределённую блокировку Redlock на Redis с Lua-скриптами;
- проектировать схемы в MongoDB, создавать индексы, анализировать планы запросов через explain();
- настраивать шардирование MongoDB, запускать change streams и применять lazy migration;
- создавать таблицы в Cassandra с правильным partition key, вставлять и читать данные через CQL;
- измерять влияние уровней консистентности (ONE, QUORUM) на задержки, запускать repair и управлять compaction;
- импортировать данные в ClickHouse, строить партиционированные таблицы, создавать материализованные представления и оптимизировать запросы;
- генерировать presigned URL в MinIO, настраивать версионирование и lifecycle policies;
- разворачивать кластер Kafka (KRaft), создавать топик с репликацией;
- реализовывать продюсеров с idempotence и acks=all, консьюмеров с ручным коммитом, демонстрировать отказоустойчивость и измерять consumer lag;
- использовать pgvector для векторного поиска (k-NN) и сравнивать производительность индексов.

владеть:

- инструментами командной строки: redis-cli, mongosh, cqlsh, clickhouse-client, mc (MinIO Client), kafka-topics, kafka-console-producer/consumer;
- техниками анализа производительности: explain () в MongoDB, nodetool в Cassandra, EXPLAIN PIPELINE в ClickHouse, consumer lag в Kafka;
- навыками локального развёртывания и тестирования распределённых систем через Docker Compose;
- методами экспериментирования с консистентностью (замеры времени при разных consistency) и персистентностью (сравнение RDB/AOF);
- подходами к эволюции схем данных в NoSQL (lazy migration, ADD/DROP колонок, Schema Registry);
- приёмами построения отказоустойчивых и восстанавливаемых кластеров (repair, репликация, exactly-once, mirroring).

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-1.	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1.	Знает основные принципы системного подхода и методы критического анализа, а также теоретические основы стратегического планирования и принятия решений
		УК-1.2.	Умеет применять методы системного анализа для выявления ключевых проблемных ситуаций, формулировать гипотезы и разрабатывать альтернативные стратегии действий на основе полученных данных
		УК-1.3.	Имеет практический опыт в проведении анализа реальных проблемных ситуаций в рамках проектов, способен вырабатывать и обосновывать стратегии действий, учитывая различные аспекты и последствия
ОПК-2.	Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.1.	Знает основные математические модели и методы, используемые в естественных науках, включая статистическое моделирование, дифференциальные уравнения и численные методы, а также современные подходы к исследованию и анализу данных в области программной инженерии в ретейле
		ОПК-2.2.	Умеет разрабатывать и адаптировать математические модели для решения конкретных проблем в естественных науках, проводить их анализ и верификацию, а также интерпретировать полученные результаты в контексте научных исследований
		ОПК-2.3.	Имеет практический опыт создания и исследования математических моделей в рамках прикладных научных

			или инновационных проектов в сфере ретейла, включая участие в публикациях, конференциях или коллаборациях, где были разработаны и апробированы новые концепции и методы
ПК-2.	Способен разрабатывать, тестировать и внедрять программные решения для цифровых систем ретейла	ПК-2.1.	Знает современные языки программирования, методы и стандарты разработки программного обеспечения, принципы проектирования и построения архитектуры информационных систем, а также подходы к обеспечению качества, надежности и безопасности программных решений в цифровых системах ретейла
		ПК-2.2.	Умеет разрабатывать алгоритмы и программные модули, интегрировать их в существующие информационные системы, проводить тестирование, отладку и сопровождение решений
		ПК-2.3.	Имеет практический опыт командной разработки и внедрения программных решений в индустриальной среде
ПК-6.	Способен представлять результаты профессиональной деятельности и формулировать практические рекомендации по развитию цифровых решений в ретейле	ПК-6.1.	Знает требования к подготовке технической и аналитической документации, принципы визуализации данных и стандарты профессиональной коммуникации
		ПК-6.2.	Умеет формулировать обоснованные выводы и рекомендации по результатам разработки и анализа, адаптируя их для технической и управленческой аудитории
		ПК-6.3.	Имеет практический опыт подготовки отчетов, презентаций и публичной защиты результатов проектов в области программной инженерии в ретейле

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Введение	3	3		14	Домашние задания
2	Redis	4	4		15	Домашние задания
3	MongoDB	4	4	4	15	Коллоквиум
4	Cassandra	4	4		15	Домашние задания
5	ElasticSearch	3	3		14	Домашние задания
6	ClickHouse	2	2	4	14	Коллоквиум
7	Объектные хранилища	2	2		13	Домашние задания
8	Kafka	2	2		13	Домашние задания
9	Векторные базы данных	2	2		13	Домашние задания
	Зачет с оценкой			4		
Итого:		26	26	12	126	
Объем дисциплины (модуля) (в ак. ч.)		190				
Объем дисциплины (модуля) (в зач. ед.)		5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение	Расширенная таксономия хранилищ. CAP, PACELC
2	Redis	Архитектура, структуры данных, персистентность Кластеризация, отказоустойчивость, оптимизация памяти
3	MongoDB	Модель данных, движок хранения, индексы Репликация, шардирование, эволюция схем
4	Cassandra	LSM-архитектура, модель данных, запись и чтение Консистентность, отказоустойчивость, compaction, эволюция схем
5	ElasticSearch	Инвертированный индекс, поиск и агрегации, распределённая архитектура
6	ClickHouse	Колоночная аналитика, внутреннее устройство, оптимизация
7	Объектные хранилища	Объектные хранилища (S3)
8	Kafka	Log-structured архитектура, продюсеры, консьюмеры Exactly-once, Kafka Streams, отказоустойчивость, мониторинг
9	Векторные базы данных	Векторные базы данных (pgvector, Milvus)

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Уильямс, Э. С++. Практика многопоточного программирования : практическое руководство / Э. Уильямс. - Санкт-Петербург : Питер, 2020. - 640 с. - (Серия «Для профессионалов»). - ISBN 978-5-4461-0831-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1733519>.

2. Цукалос, М. Golang для профи: работа с сетью, многопоточность, структуры данных и машинное обучение с Go : практическое руководство / М. Цукалос. - Санкт-Петербург : Питер, 2021. - 720 с. - (Серия «Для профессионалов»). - ISBN 978-5-4461-1617-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1733707>.

Дополнительная литература:

1. Бэнкер, К. MongoDB в действии : практическое руководство / К. Бэнкер ; пер. с англ. А. А. Слинкина. - 2-е изд. - Москва : ДМК Пресс, 2023. - 396 с. - ISBN 978-5-89818-344-8.

2. Скотт, Д. Kafka в действии : практическое руководство / Д. Скотт, В. Гамов, Д. Клейн ; пер. с англ. А. Н. Киселева. - Москва : ДМК Пресс, 2022. - 310 с. - ISBN 978-5-93700-118-4.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья,

оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		

Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Системы управления базами данных» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, коллоквиумы, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал, использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Коллоквиум – устные ответы на вопросы, список которых известен студенту заранее.

В процессе подготовки к коллоквиуму необходимо проанализировать учебные материалы.

материалы, ознакомившись с лекциями, учебниками и дополнительными источниками, акцентируя внимание на ключевых темах. Рекомендуется создать структурированные конспекты, выделяя основные идеи, термины и формулы.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Системы управления базами данных»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме *зачета с оценкой*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	
8	Отлично	
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать
6	Хорошо	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Системы управления базами данных» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	40%	6	Набор задач по темам недели
Коллоквиум	40%	2	Устные ответы на вопросы, список которых известен студенту заранее
Зачет с оценкой	20%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Системы управления базами данных»: $0,4 \times \text{среднее за домашние задания} + 0,4 \times \text{среднее за коллоквиумы} + 0,2 \times \text{зачет с оценкой}$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1

1. Разверните Redis с помощью Docker, настройте персистентность (RDB и AOF). Сравните время восстановления при сбое.
2. Реализуйте rate limiter для API (10 запросов в минуту на IP) с использованием Redis и Lua-скрипта.
3. Создайте очередь задач с приоритетами с помощью Redis (используя sorted set).
4. Настройте Redis Sentinel для отказоустойчивости. Протестируйте failover.
5. Разверните кластер Redis (6 узлов) с хэш-слотами. Проверьте распределение данных.

Домашнее задание 2

1. Установите MongoDB через Docker, создайте коллекцию пользователей с вложенными полями.
2. Постройте составной индекс и проанализируйте план запроса с помощью **explain()**.
3. Настройте реплика-сет из 3 узлов. Протестируйте поведение при падении primary.
4. Реализуйте change stream для отслеживания изменений в коллекции.
5. Выполните lazy migration: добавьте новое поле, обновите данные пачками, переключите приложение.

Домашнее задание 3

1. Разверните кластер Cassandra (3 узла) через Docker Compose. Создайте таблицу с правильным partition key.
2. Вставьте 10 000 строк, измерьте задержки при чтении с уровнями консистентности ONE и QUORUM.
3. Запустите **nodetool repair**, объясните, когда это необходимо.
4. Создайте топик в Kafka (KRaft), настройте репликацию.
5. Напишите продюсера с **acks=all** и идемпотентностью, консьюмера с ручным коммитом. Измерьте consumer lag.

Примерные вопросы к коллоквиуму

Занятие 1:

1. В чём разница между RDB и AOF в Redis? Какой метод быстрее при восстановлении?
2. Что такое Redlock? Какие проблемы он решает и какие у него ограничения?
3. Как устроен WiredTiger в MongoDB? Как работает MVCC?
4. Что такое oplog? Как он используется при репликации?
5. Как работает шардирование в MongoDB? Что такое shard key и balancer?
6. Что такое LSM-дерево? Какие компоненты входят в его архитектуру?
7. Как Cassandra обрабатывает запрос на чтение при QUORUM?
8. Что такое tombstone в Cassandra? Почему его важно удалять?
9. Как работает hinted handoff? В каких случаях используется?
10. Что такое compaction в Cassandra? Какие стратегии существуют?

Занятие 2:

1. Чем колоночное хранение отличается от строкового? В чём выгода для аналитики?
2. Что такое MergeTree в ClickHouse? Как влияют ORDER BY и PRIMARY KEY?
3. Как работает векторизация в ClickHouse? Приведите пример ускорения запроса.
4. Что такое presigned URL в S3? Зачем он нужен?
5. Как устроена модель данных в Kafka? Что такое партиция и оффсет?
6. Как Kafka обеспечивает exactly-once семантику?
7. Что такое consumer group? Как Kafka распределяет сообщения между потребителями?
8. Как работает инвертированный индекс в Elasticsearch?
9. Какие метрики близости используются в векторных БД? Приведите примеры.
10. В чём разница между HNSW и IVF? Какой индекс быстрее при поиске?

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Как называется теорема, описывающая компромисс между согласованностью, доступностью и устойчивостью к разбиению сети?	CAP	УК-1
2.	Какая модель данных используется в MongoDB?	документная	ПК-2
3.	Какой тип персистентности в Redis основан на снимках памяти?	RDB	ОПК-2
4.	Какая структура данных в Redis используется для реализации очереди с приоритетами?	sorted set	ПК-6

5.	Какой движок хранения используется по умолчанию в MongoDB?	WiredTiger	ПК-2
6.	Как называется журнал операций, используемый для репликации в MongoDB?	oplog	ОПК-2
7.	Какая архитектура хранения лежит в основе Cassandra?	LSM-дерево	УК-1
8.	Какой фильтр используется в Cassandra для быстрой проверки наличия ключа в SSTable?	Bloom-фильтр	ПК-6
9.	Какой уровень консистентности в Cassandra требует ответа от большинства реплик?	QUORUM	ПК-2
10.	Как называется протокол обмена состоянием между узлами в Cassandra?	gossip	ОПК-2
11.	Как называется процесс слияния SSTable в Cassandra для освобождения места и ускорения чтения?	compaction	ПК-2
12.	Какой движок таблиц используется по умолчанию в ClickHouse?	MergeTree	ПК-6
13.	Как называется тип хранения данных, при котором значения хранятся по столбцам?	колоночное	УК-1
14.	Как называется логический контейнер для объектов в S3?	bucket	ОПК-2
15.	Как называется единица хранения в S3, содержащая данные и метаданные?	object	ПК-6
16.	Как называется временная ссылка с токеном доступа для чтения объекта в S3?	presigned URL	ПК-2
17.	Какая архитектура Kafka основана на последовательной записи логов?	log-structured	ОПК-2
18.	Какая гарантия доставки сообщений в Kafka обеспечивает ровно одно получение?	exactly-once	ПК-6
19.	Какой индекс в векторных БД построен на графе ближайших соседей?	HNSW	УК-1
20.	Как называется расширение PostgreSQL для хранения и поиска векторов?	pgvector	ПК-2