



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(ДГТУ)**

Факультет
Информатика и вычислительная техника

Кафедра
Программное обеспечение вычислительной техники и автоматизированных систем

УСТАНОВКА И НАПОЛНЕНИЕ РАСПРЕДЕЛЕННОЙ БАЗЫ ДАННЫХ

Практикум

по выполнению лабораторной работы №4
по дисциплине «Распределенные системы обработки и хранения информации»

Технология разработки сложных программных систем
(указывается направленность (профиль) образовательной программы)

09.04.04 Программная инженерия
(указывается код и наименование направления подготовки)

Ростов-на-Дону
2024 г.

Составители: канд. техн. наук, зав. каф. В.В. Долгов

УДК 004.75

Базовые типы данных, управляющие конструкции и функции: метод. указания. – Ростов н/Д: Издательский центр ДГТУ, 2024.

В методическом указании рассматриваются вопросы установки и использования распределенных баз данных. Приведены задания к лабораторной работе, помогающие закрепить на практике полученные знания, и контрольные вопросы для самопроверки. Предназначено для обучающихся по направлению 09.04.04 «Программная инженерия», профиль «Технология разработки сложных программных систем».

Ответственный за выпуск:

зав. кафедрой «Программное обеспечение вычислительной техники и автоматизированных систем» Долгов В.В.

© В.В. Долгов, 2024

© Издательский центр ДГТУ, 2024

1 Теоретическая часть

2 Задание к лабораторной работе

В соответствие с заданием из таблицы 1 выполнить развертывание и заполнение указанной базы данных данными. После заполнения базы данных, проверить корректность её работы и доступность данных на ряде простых аналитических запросов.

Таблица 1. Варианты заданий

| № варианта | Задание к лабораторной работе |
|------------|---|
| 1 | База данных: Apache Cassandra. 1. Разверните кластер Apache Cassandra в трех docker-контейнерах, настроенных для совместной работы. 2. Соберите собственный docker-контейнер с программой, которая генерирует большой объем данных (не менее 100 000 записей) о транзакциях интернет-магазина (пользователь, товар, цена, дата покупки) и сохраняет их в Cassandra. 3. Выполните запросы для получения общей суммы продаж за определенный период. 4. Найдите наиболее популярные товары. 5. Проверьте распределение покупок по регионам. |
| 2 | База данных: OpenSearch. 1. Установите кластер OpenSearch из трех узлов в docker-контейнерах. 2. Создайте docker-контейнер с программой, которая индексирует большой набор текстовых данных (например, коллекцию новостных статей) в OpenSearch. 3. Реализуйте полнотекстовый поиск по ключевым словам. 4. Найдите статьи по диапазону дат. 5. Проведите агрегацию данных для определения наиболее часто встречающихся тем. |
| 3 | База данных: Apache HBase. 1. Разверните HBase поверх Hadoop в docker-контейнерах, настроив распределенную файловую систему HDFS. 2. Соберите docker-контейнер с программой, которая обрабатывает большой объем логов серверов и сохраняет их в HBase для дальнейшего анализа. 3. Выведите логи за определенный период. 4. Найдите в логах ошибки серверов заданного типа. 5. Проанализируйте частоту возникновения ошибок по времени. |
| 4 | База данных: Neo4j. 1. Установите Neo4j в docker-контейнерах и настройте кластерный режим работы. |

| | |
|---|---|
| | <p>2. Создайте docker-контейнер с программой, которая строит социальный граф на основе больших данных о связях между людьми (друзья, коллеги, родственники) и сохраняет его в Neo4j.</p> <p>3. Найдите кратчайший путь связи между двумя людьми.</p> <p>4. Определите наиболее влиятельных участников сети.</p> <p>5. Проанализируйте сообщества внутри графа.</p> |
| 5 | <p>База данных: InfluxDB.</p> <p>1. Разверните InfluxDB в docker-контейнерах с настройкой кластеризации для горизонтального масштабирования.</p> <p>2. Соберите docker-контейнер с программой, которая собирает и сохраняет телеметрические данные с большого количества IoT-устройств (например, датчики температуры, влажности).</p> <p>3. Постройте график изменения температуры за последний месяц.</p> <p>4. Найдите устройства с аномальными показателями.</p> <p>5. Проанализируйте корреляцию между различными сенсорами.</p> |
| 6 | <p>База данных: ClickHouse.</p> <p>1. Установите кластер ClickHouse из нескольких узлов в docker-контейнерах для обработки аналитических запросов.</p> <p>2. Создайте docker-контейнер с программой, которая загружает большой объем данных веб-аналитики (посещения страниц, время на сайте, переходы) в ClickHouse.</p> <p>3. Подсчитайте общее количество уникальных посетителей за день.</p> <p>4. Определите наиболее посещаемые страницы.</p> <p>5. Проанализируйте пути переходов пользователей по сайту.</p> |

3 Материально-техническое обеспечение работы

Аудитория для проведения лабораторных занятий должна быть укомплектована специализированной мебелью и индивидуальными компьютерами следующей минимальной комплектации:

- Процессор: не менее двух исполнительных ядер, совместимый с системой команд x86 и x64, с поддержкой аппаратной виртуализации.
- Оперативная память: не менее 8 Гб.
- Монитор: не менее 24" (дюймов) по диагонали.
- Наличие локальной сети со скоростью обмена не менее 1 Гб/сек.
- Наличие доступа в сеть Интернет со скоростью не менее 1 Мбит/сек.
- Наличие клавиатуры и манипулятора «мышь».

На компьютерах должно быть установлено следующее программное обеспечение:

- Операционная система: любая современная операционная система не ранее 2022 года выпуска допускающая установку JDK версии не ниже 21 и среды разработки программного обеспечения, поддерживающей язык программирования Kotlin.
- Среда разработки: IntelliJ IDEA Community Edition или Visual Studio Code.
- Среда исполнения: JDK версии не ниже 21.

4 Порядок выполнения и сдачи работы

Для выполнения лабораторной работы рекомендуется придерживаться следующего порядка выполнения:

1. Ознакомиться с темой и целями лабораторной работы
2. Изучить теоретический материал
3. Подготовить рабочее окружение
4. Разработать программный код лабораторной работы в соответствии с заданием
5. Произвести проверку работоспособности, выполнить тестирование и отладку кода
6. Проанализировать полученные результаты и примененные в ходе решения подходы
7. Выполнить самооценку и рефлексию
8. Сдать лабораторную работу преподавателю и получить от него обратную связь

При сдаче студентом лабораторной работы основным отчетом выступает исходный код самостоятельно созданной в процессе выполнения работы программы.

Исходный код должен быть отформатирован согласно принятым для используемого языка программирования стандартам. Является желательным наличие в исходном коде комментариев, описывающих основные части программы и особенности их функционирования. В то же время студент должен быть готов объяснить работу программы в целом и каждую отдельную ее часть при полном отсутствии комментариев (например, они могут быть удалены преподавателем при сдаче работы).

Обязательным условием сдачи является умение студента самостоятельно восстановить любой участок исходного кода программы (но не более 20 строк подряд) после его удаления. Удаленный участок должен быть самостоятельно восстановлен

обучающимся заново в присутствии преподавателя. При восстановлении удаленного участка кода запрещается использовать операции Undo (Отменить) текстовых редакторов или переписывание кода участка из других источников.

Исходный код сдаваемой программы должен быть представлен в электронном виде.

Сдача исходного кода ранее сдававшихся программ или программ, код которых выложен в сети Интернет, не допускается.

5 Контрольные вопросы к лабораторной работе

1. Какие

6 Перечень использованных информационных ресурсов

1. Kotlin Docs | Kotlin Documentation [Электронный ресурс]. – URL: <https://kotlinlang.org/docs/home.html>
2. IntelliJ IDEA – the Leading Java and Kotlin IDE [Электронный ресурс]. – URL: <https://www.jetbrains.com/idea/>
3. Visual Studio Code - Code Editing. Redefined [Электронный ресурс]. – URL: <https://code.visualstudio.com/>

Редактор А.А. Литвинова

| | | |
|--------------------------------|---------|------------------|
| ЛР № 04779 от 18.05.01. | В набор | В печать |
| Объем 0,5 усл.п.л., уч.-изд.л. | Офсет. | Формат 60x84/16. |
| Бумага тип №3. | Заказ № | Тираж 75. Цена |

Издательский центр ДГТУ

Адрес университета и полиграфического предприятия:

344010, г. Ростов-на-Дону, пл. Гагарина, 1.