

Департамент образования и науки города Москвы
Государственное автономное образовательное учреждение
высшего образования города Москвы
«Московский городской педагогический университет»
Институт цифрового образования
Департамент информатики управления и технологий

Кузьмина Дарья Юрьевна БД-241м

Инструменты хранения и анализа больших данных

Лабораторная работа 2.1. Часть 2.
Cassandra
Вариант 11

Направление подготовки/специальность
38.04.05 - Бизнес-информатика
Бизнес-аналитика и большие данные
(очная форма обучения)

Руководитель дисциплины:
Босенко Т.М., доцент департамента
информатики, управления и технологий,
доктор экономических наук

Москва
2025

Содержание

Введение	2
Основная часть	2
Заключение	7

Введение

Цель:

Получить практические навыки работы с базой данных Cassandra, изучив основные операции по управлению данными, включая создание и использование ключспейсов, таблиц, выполнение запросов CQL, а также работу с различными инструментами подключения и администрирования.

Задачи:

1. Подключиться к Cassandra через cqlsh или браузерные интерфейсы (Cassandra Web, Apache Zeppelin).
2. Создать ключспейс с заданной стратегией репликации (SimpleStrategy).
3. Создать таблицы для хранения данных (например, movies, actors) с использованием примитивных типов и коллекций.
4. Выполнить операции CRUD (добавление, выборка, обновление, удаление данных) в созданных таблицах.
5. Изучить метаданные ключспейсов и таблиц с помощью команд DESCRIBE и запросов к системным таблицам.

Основная часть

Задача 1.

В виртуальной машине проверяем, что сервер запущен:

```
● dev@dev-vm: ~/Downloads/dba/nonrel/mongo$ cd ..  
● dev@dev-vm: ~/Downloads/dba/nonrel$ cd cassandra
```

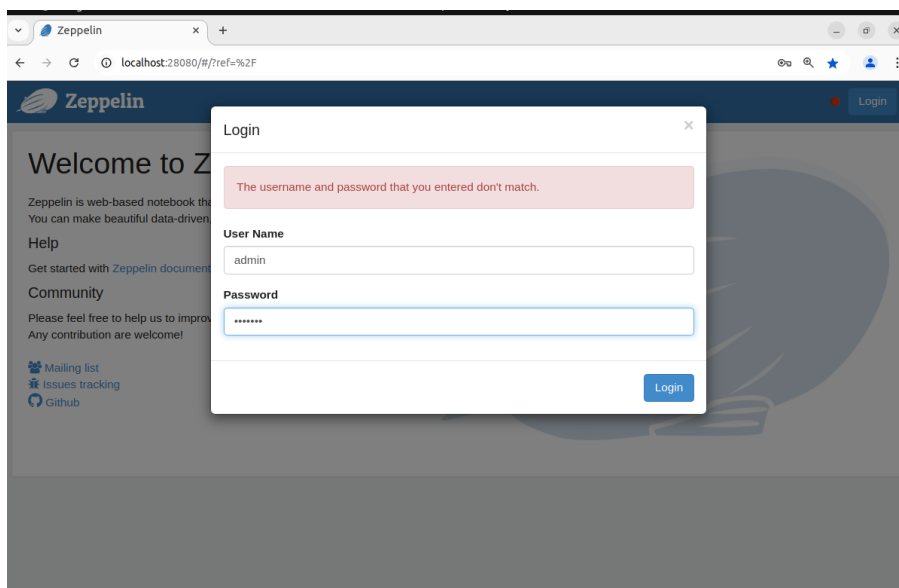
```
● dev@dev-vm: ~/Downloads/dba/nonrel/cassandra$ docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	P
b7804eb56d	mongo:7.0.17-rc1-jammy	"docker-entrypoint.s..."	56 minutes ago	Up 56 minutes	0
.0.0.0:27017->27017/tcp, [::]:27017->27017/tcp	mongo-1				
a5bb7206afbd	postgres:16	"docker-entrypoint.s..."	6 weeks ago	Up About an hour	0
.0.0.0:5432->5432/tcp, [::]:5432->5432/tcp	postgres16				
ec4919b750da	dpape/pgadmin4:latest	"/entrypoint.sh"	6 weeks ago	Up About an hour	0
.0.0.0:80->80/tcp, [::]:80->80/tcp, 443/tcp	pgadmin				

```
○ dev@dev-vm: ~/Downloads/dba/nonrel/cassandra$
```

Убеждаемся, что запущены 3 контейнера

Заходим в Zeppelin, пароль не подходит.



Индивидуальное задание **Вариант 11**

1. Создайте ключспейс shops с репликацией SimpleStrategy и коэффициентом репликации 1.
2. Создайте таблицу products в ключспейсе shops с полями product_id (int), name (text), price (float), category (text), stock (int) и первичным ключом product_id.
3. Вставьте четыре продукта в таблицу products.
4. Выберите все продукты из таблицы products.
5. Обновите поле stock продукта с product_id = 2.

Добавляем свои данные с гитхаба.

Судя по лекции, концептуально, это задание выглядело бы так:

1. Создание ключспейса shops

Для начала работы необходимо создать ключспейс (пространство ключей) с именем "shops", используя стратегию репликации SimpleStrategy и коэффициент репликации 1 (так как мы работаем с одним узлом Cassandra):

```
CREATE KEYSPACE shops WITH replication = {  
    'class': 'SimpleStrategy',  
    'replication_factor': 1  
};
```

Эта команда создает новое пространство ключей, где:

- 'SimpleStrategy' - простая стратегия репликации, подходящая для одноподатацентровых кластеров
- 'replication_factor: 1' - указывает, что каждая запись будет храниться только на одном узле

2. Создание таблицы products

После создания ключспейса переходим в него и создаем таблицу products:

USE shops;

```
CREATE TABLE products (  
    product_id int,  
    name text,  
    price float,  
    category text,  
    stock int,  
    PRIMARY KEY (product_id)  
);
```

Таблица содержит следующие поля:

- `product_id` - уникальный идентификатор продукта (первичный ключ)
- `name` - название продукта
- `price` - цена продукта
- `category` - категория продукта
- `stock` - количество товара на складе

3. Вставка данных в таблицу products

Добавим четыре продукта в таблицу:

```
INSERT INTO products (product_id, name, price, category, stock)  
VALUES (1, 'Laptop', 999.99, 'Electronics', 15);
```

```
INSERT INTO products (product_id, name, price, category, stock)  
VALUES (2, 'Smartphone', 699.99, 'Electronics', 25);
```

```
INSERT INTO products (product_id, name, price, category, stock)
VALUES (3, 'Headphones', 149.99, 'Accessories', 50);
```

```
INSERT INTO products (product_id, name, price, category, stock)
VALUES (4, 'Keyboard', 49.99, 'Accessories', 30);
```

4. Выборка всех продуктов

Проверим, что данные успешно добавлены, выполнив запрос на выборку всех записей:

```
SELECT * FROM products;
```

Ожидаемый результат:

product_id	name	price	category	stock
1	Laptop	999.99	Electronics	15
2	Smartphone	699.99	Electronics	25
3	Headphones	149.99	Accessories	50
4	Keyboard	49.99	Accessories	30

5. Обновление количества товара на складе

Обновим количество товара на складе для продукта с product_id = 2:

```
UPDATE products SET stock = 30 WHERE product_id = 2;
```

После выполнения этой команды количество товара "Smartphone" изменится с 25 до 30.

Убедимся, что обновление прошло успешно:

```
SELECT * FROM products WHERE product_id = 2;
```

Ожидаемый результат:

product_id	name	price	category	stock
2	Smartphone	699.99	Electronics	30

Заключение

В теории удалось получить практические навыки работы с базой данных Cassandra, изучив основные операции по управлению данными, включая создание и использование ключспейсов, таблиц, выполнение запросов CQL, а также работу с различными инструментами подключения и администрирования.