

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО”**

Факультет прикладної математики
Кафедра програмного забезпечення комп’ютерних систем

КУРСОВА РОБОТА

з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: “Система аналізу цін споживчих товарів”

Студент

групи КП-91(92,93)

Ісачко І. В.

(підпис)

Викладач

к.т.н, доцент кафедри

СПіСКС

Петрашенко А.В.

(підпис)

Захищено з оцінкою _____

Київ – 2021

Анотація

“Система аналізу цін споживчих товарів”. Курсова робота полягає у закріпленні навичок та знань за допомогою спеціально визначених практичних завдань курсу. Зміст роботи полягає у розробленні програмного додатку. Додаток має консольний вигляд, але для графічної та серверної частини були використані додаткові ресурси. Програма реалізує поступові етапи для розвитку та вдосконалення більш вразливих частин.

Для генерування адекватних даних використовується інформаційні ресурси (інтернет магазини споживчих товарів). Реалізована валідація дозволяє уникнути небажаних повторів однорідної інформації в системі цін й оновити застарілу. Щоб зберігати та обробляти згенеровану інформацію було спроектовано базу даних, яка відповідає третій нормальній формі. Зазвичай стаються випадки коли важлива інформація безслідно зникає разом з сервером, де вона зберігалась. Для уникнення негативних наслідків, які спричиняє деактивація сервера, дані реплікуються в автоматичному режимі на іншому сервері. Це дозволяє уникнути жахливих ситуацій й швидко повернути базу даних до життя.

Переглянути файли роботи можна на GitHub репозиторію, за посиланням https://github.com/luamoris/db_coursework.

Зміст

Аналіз інструментарію для виконання курсової роботи	5
Опис програмного забезпечення	6
Аналіз функціонування засобів реплікації	8
Аналіз функціонування засобів резервування/відновлення	9
Аналіз результатів підвищення швидкодії запитів	10
Опис результатів аналізу предметної галузі	11

Вступ

Призначення роботи полягає у створенні надійного сервера, який зберігає та обробляє збережену інформацію. Крім того роботу забезпечують надійні консольний та графічний додатки для зручності використання всіх реалізованих модулів.

Актуальність продажі товару у світі досить складно переоцінити. Надзвичайно потужна економічна система у світі ніколи не перестане існувати, адже пока є попит будуть і пропозиції. Споживчі товари займають що не менш призове місце в житті кожної, без винятків (окрім можливо лише ще нецивілізованих людей, але й в них є щось подібне). Отже актуальність теми не можливо не помітити. Й мені вдалося, на мій погляд, як помітити, так і реалізувати інформаційно-аналітичну систему широкого призначення, яка дозволяє отримувати інформацію або генерувати та здійснювати її обробку та аналіз.

Загальна постановка завдання складається з таких пунктів:

1. Реалізація системи для генерації адекватних даних предметної галузі.
2. Реалізація засобів фільтрації та валідації згенерованих даних для знаходження більш корисних.
3. Розроблення бази даних призначеної для зберігання, аналізу та реплікації інформації.
4. Забезпечення засобів реплікації даних, що зберігаються у СУБД.
5. Реалізація алгоритмів аналізу даних або використання методів машинного навчання.
6. Забезпечення засобів резервування та відновлення даних.

Аналіз інструментарію для виконання курсової роботи

СУБД PostgreSQL — головний ресурс, який зі зручним інтерфейсом користувача дозволяє зберігати дані, та робити будь-які маніпуляції над ними. Крім того додаткові можливості, як інтегрування мов програмування в інструментарій, тригерні функції, представлення та багато іншого роблять цю СУБД унікальною.

Web scraping - вільні технології, разом з бібліотеками, такі як cheerio дають змогу використовувати дані зі сторінок інтернету без використання API.

За допомогою бібліотеки d3 технології JavaScript був реалізований графічний візуал у веб-додатку. Інформації щодо використання саме цієї бібліотеки небагато. Але існують талановиті програмісти, які не шкодують своїх розробок для вивчення іншим.

Express модулі технології NodeJs дозволили запустити власний сервер, який забезпечує доступ до алгоритмів перегляду та аналізу інформації по темі курсової роботи.

Здебільш за допомогою інтернет ресурсів вдалося реалізувати реплікацію даних на двох серверах (де перший master, другий — slave). Проте більш значні знання вдалося закріпити на курсі “OS Linux”. З вивченого курсу були використані знання створення віртуальних машин, налаштування та ведення системного адміністрування.

Деяка інформація в мережі не є достовірною, та призводить до проблем, тому потрібно якісно вивчати питання перед його вирішенням.

Опис програмного забезпечення

Курсова робота складається з трьох структурних частин, які працюють як різні додатки, але пов'язані з єдиними ресурсами бази даних.

Генерація даних (web scraping).

- `\scraping\rozetka.js`
- лише один файл (скрипт), який виконує функцію збереження даних з інтернет магазину розетка.

Як консольний так й графічний додатки використовують розроблений спільний модуль, який взаємодіє з базою даних, та власне сервером.

- `\database\models\product.js`
- модель, яка представляє об'єкт “Продукт”.
- `\database\models\category.js`
- модель, яка представляє об'єкт “Категорія”.
- `\database\models\pgsql.js`
- модуль, який забезпечує взаємозв'язок між клієнтом (розробником) та сервером з даними.
- `\database\storage\db.js`
- модуль, який реалізує більш широкий спектр дій для роботи з даними, використовуючи модуль нижнього рівня `pgsql`.

Консольний додаток.

- дозволяє за допомогою певних команд, введених в консоль, швидко знаходити потрібну інформацію з сервера даних.
- `\routers\router.js` й `\controllers\controller.js`
- реалізує модель роутер — контроллер.

- `\view\show.js`
- модуль, який створює функцію для початку роботи у консолі.
- **app.js**
- головний файл, який запускає роботу консольного додатку.

Графічний додаток.

- додаток забезпечує графічну взаємодію користувача з сервером даних, для зображення інформації у графіках.
- `\www\site.js`
- експрес сервер, який забезпечує передачу даних до технології frontend, за допомогою модулю Ajax.
- модуль обробляє запити та віддає дані, які відповідають запиту.
- `\www\views\index.html`
- головний файл з візуальним представленням графічного додатку.
- `\www\public\script.js`
- забезпечує отримання даних з сервера, їх обробку та використання для створення графіків.
- `\www\public\style.css`
- стилістика веб-додатку.

Аналіз функціонування засобів реплікації

При будь яких змінах, які відбуваються з даними, на головному сервері (master) ці ж дані автоматично змінюється на підрядному сервері (slave). Дублювання на підрядний сервер відбувається при оновленні, видаленні або генерації нових даних на головному сервері. Але використати такі ж самі дії на підрядному сервері не дозволено. Такі дані можливо лише зчитувати та використовувати для пошуку.

Підрядний сервер починає приймати запити лише коли головний сервер зробився недієздатним.

Тестування проводилось шляхом відключення головного сервера.

Аналіз функціонування засобів резервування/відновлення

Дана можливість не була реалізована.

Аналіз результатів підвищення швидкодії запитів

Для вирішенні проблем з тривалими запитами були застосовані дії для їх уникнення.

По перше, на сервері з даними був створений тригер, який забезпечує вирішення проблеми дублювання однакових даних. Проте для дослідження цінової політики однакові дані з різною ціною — необхідні. Тому тригерна функція вирішує цю проблему й проблему з оптимізацією. Запит був створений саме на сервері, бо сервер швидше виконує задачу перевірки даних, чим перевірка за допомогою клієнта.

По друге, на основі частого використання деяких полів в запитах були створенні 2 індекси, які значно підвищили швидкість виконання запитів. Індекси мають однаковий тип — BTree. З тих причин, що атрибути цілочисельного значення й використовуються при пошуку для умови.

Приклади у додатках.

Опис результатів аналізу предметної галузі

Результати аналізу обраної предметної галузі стимулюють на наступні кроки у її вивченні. Динаміка цін змінюється на основні соціально значущих продовольчих товарів.

Аналіз предметної галузі можна переглянути у графічних додатках (графіки).

Висновки

Підведемо підсумки. Під час виконання курсової роботи була досягнута її мета. А саме досягнення практичних навичок розробки сучасного програмного забезпечення, що взаємодіє з реляційними базами даних, а також здобуття навичок оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації.

Головним результатом став досвід, який обов'язково буде використано у власних проектах.

Література

1. Встановити Linux Ubuntu Server 20.04. Детальна інструкція для початківців. URL
<https://info-comp.ru/install-linux-ubuntu-server-20-04>
2. Встановити і налаштування PostgreSQL 13 на Linux Ubuntu Server 20.04. URL
<https://info-comp.ru/install-postgresql-13-on-ubuntu-server>
3. How To Set Up a Firewall with UFW on Ubuntu 20.04. URL
<https://www.digitalocean.com/community/tutorials/how-to-set-up-a-firewall-with-ufw-on-ubuntu-20-04>
4. How To Set Up Physical Streaming Replication with PostgreSQL 12 on Ubuntu 20.04. URL
<https://www.digitalocean.com/community/tutorials/how-to-set-up-physical-streaming-replication-with-postgresql-12-on-ubuntu-20-04>

Додатки

categories

Общие Столбцы Дополнительно Ограничения Параметры Безопасность SQL

Наследуется из таблиц(ы)

Выберите источник нас

Столбцы

	Имя	Тип данных	Length/Precision	Масштаб	Не NULL?	Первичный ключ?
	id	integer			<input checked="" type="checkbox"/> Да	<input checked="" type="checkbox"/> Да
	category	text			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет

Рис. Таблица “categories”

products

Общие Столбцы Дополнительно Ограничения Параметры Безопасность SQL

Наследуется из таблиц(ы)

Выберите источник нас

Столбцы

	Имя	Тип данных	Length/Precision	Масштаб	Не NULL?	Первичный ключ?
	id	integer			<input checked="" type="checkbox"/> Да	<input checked="" type="checkbox"/> Да
	goods_id	integer			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет
	title	text			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет
	price	real			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет
	availability	boolean			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет
	date	timestamp without time zone			<input checked="" type="checkbox"/> Да	<input type="checkbox"/> Нет
	url	text			<input type="checkbox"/> Нет	<input type="checkbox"/> Нет
	category_id	integer			<input type="checkbox"/> Нет	<input type="checkbox"/> Нет

Рис. Таблица “products”

```

CREATE OR REPLACE FUNCTION func_product_insert()
RETURNS TRIGGER AS
$$
    DECLARE
        product products%ROWTYPE;
    BEGIN
        SELECT * INTO product FROM products WHERE goods_id = NEW.goods_id ORDER BY date DESC LIMIT 1;
        IF NOT FOUND THEN
            RETURN NEW;
        ELSE
            IF product.price = NEW.price AND NEW.date < product.date + interval '1day' THEN
                RAISE EXCEPTION 'The product is already in the table';
            ELSE
                RETURN NEW;
            END IF;
        END IF;
    END;
$$
LANGUAGE 'plpgsql';

```

Рис. Створення функції “func_product_insert()”

```

CREATE TRIGGER tg_product_insert
BEFORE INSERT
ON products
FOR EACH ROW
EXECUTE PROCEDURE func_product_insert();

```

Рис. Створення триггера “tg_product_insert”

```

CREATE INDEX index_category_id ON products USING BTREE(category_id);
CREATE INDEX index_goods_id ON products USING BTREE(goods_id);

```

Рис. Створення індексів

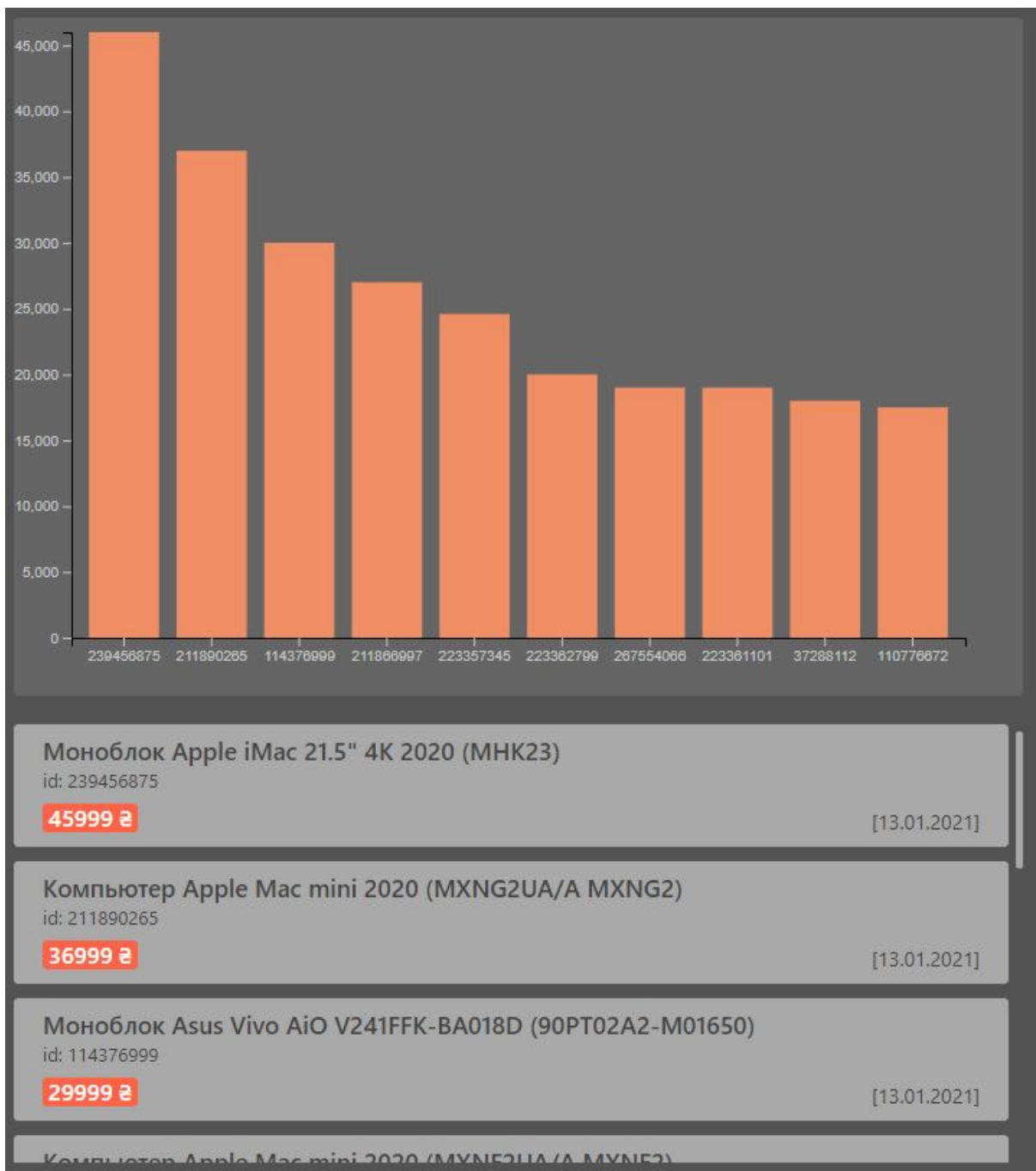
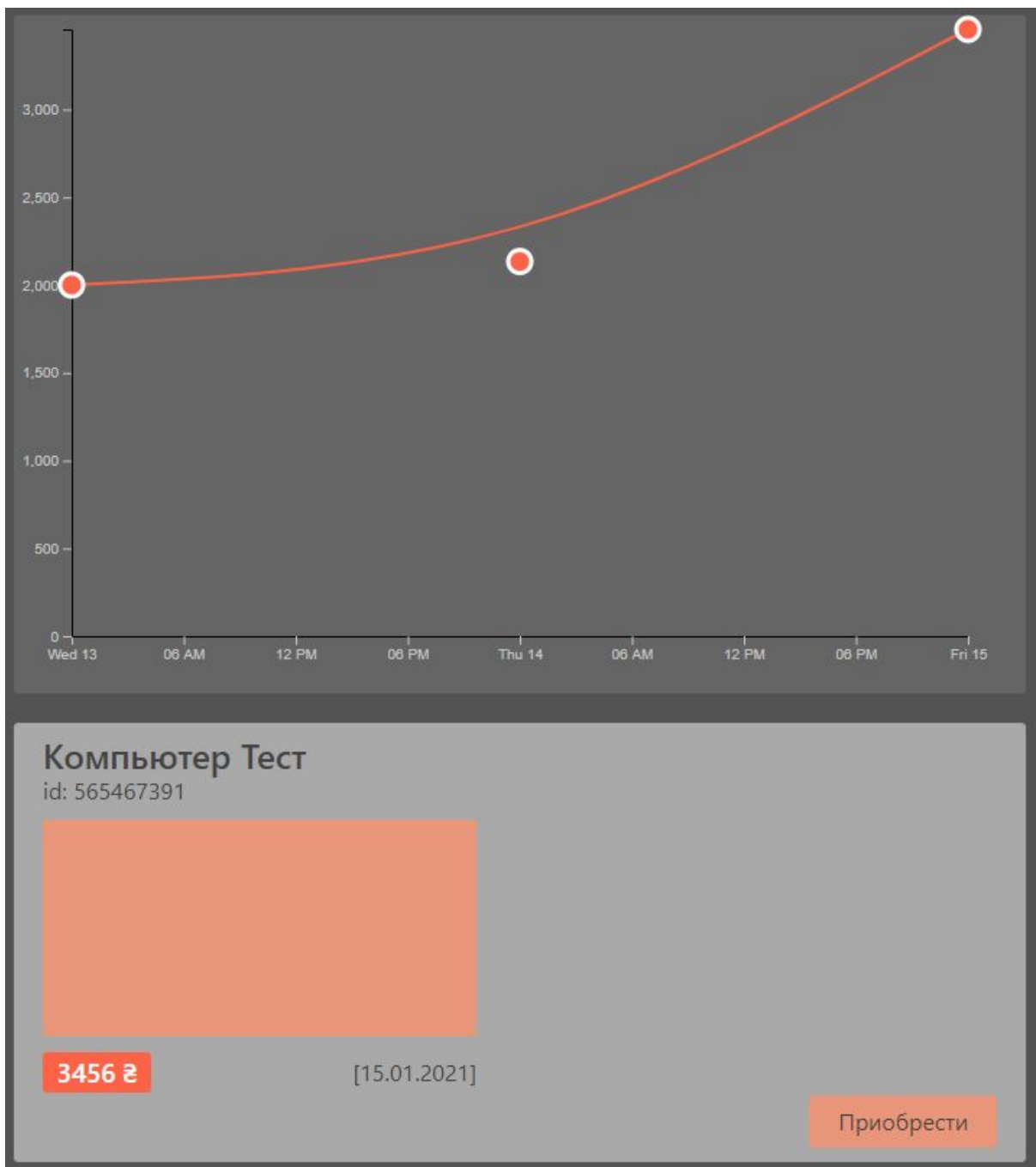


Рис. Графік 10-ти найбільш дорогих “продуктів” з категорії “комп'ютери”



***Рис.** Графік зміни ціни на “продукт” з часом*

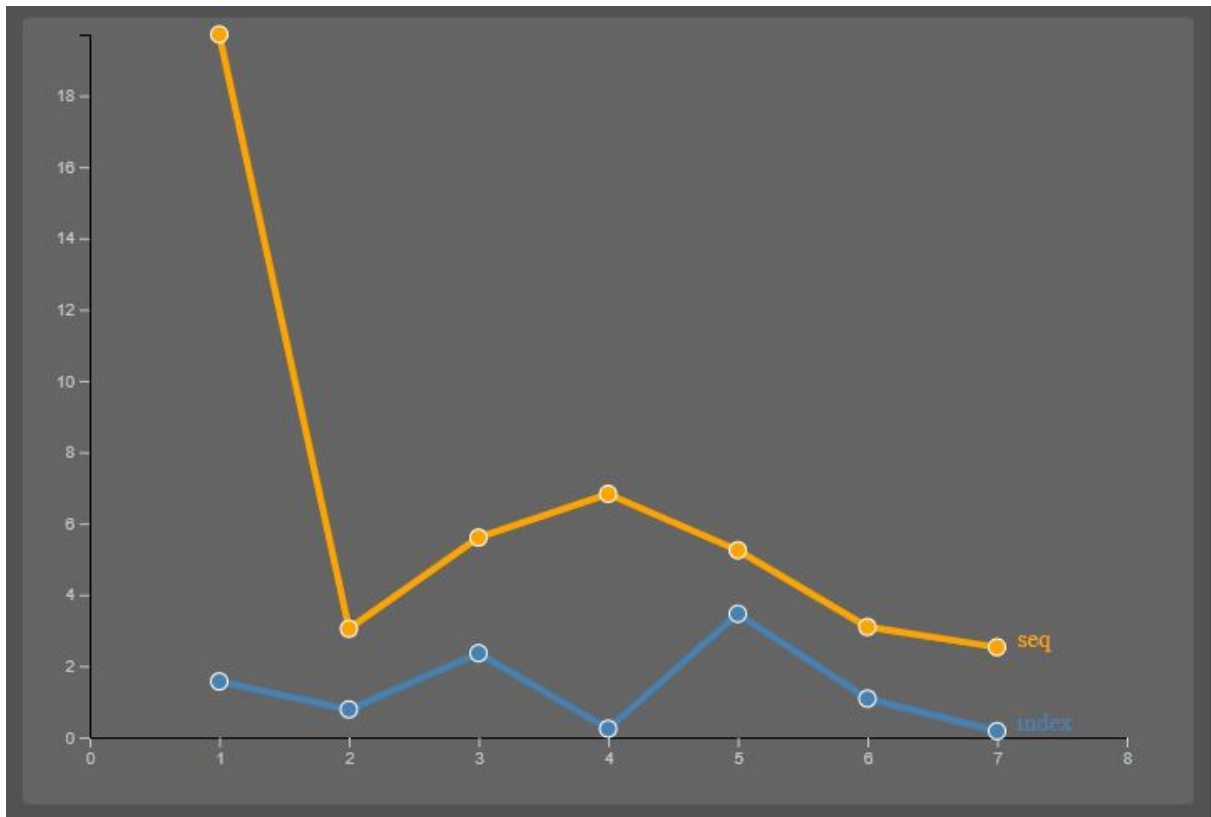


Рис. Графік швидкодії запитів з використанням *index*(блакитна) та *seq*(жовта)

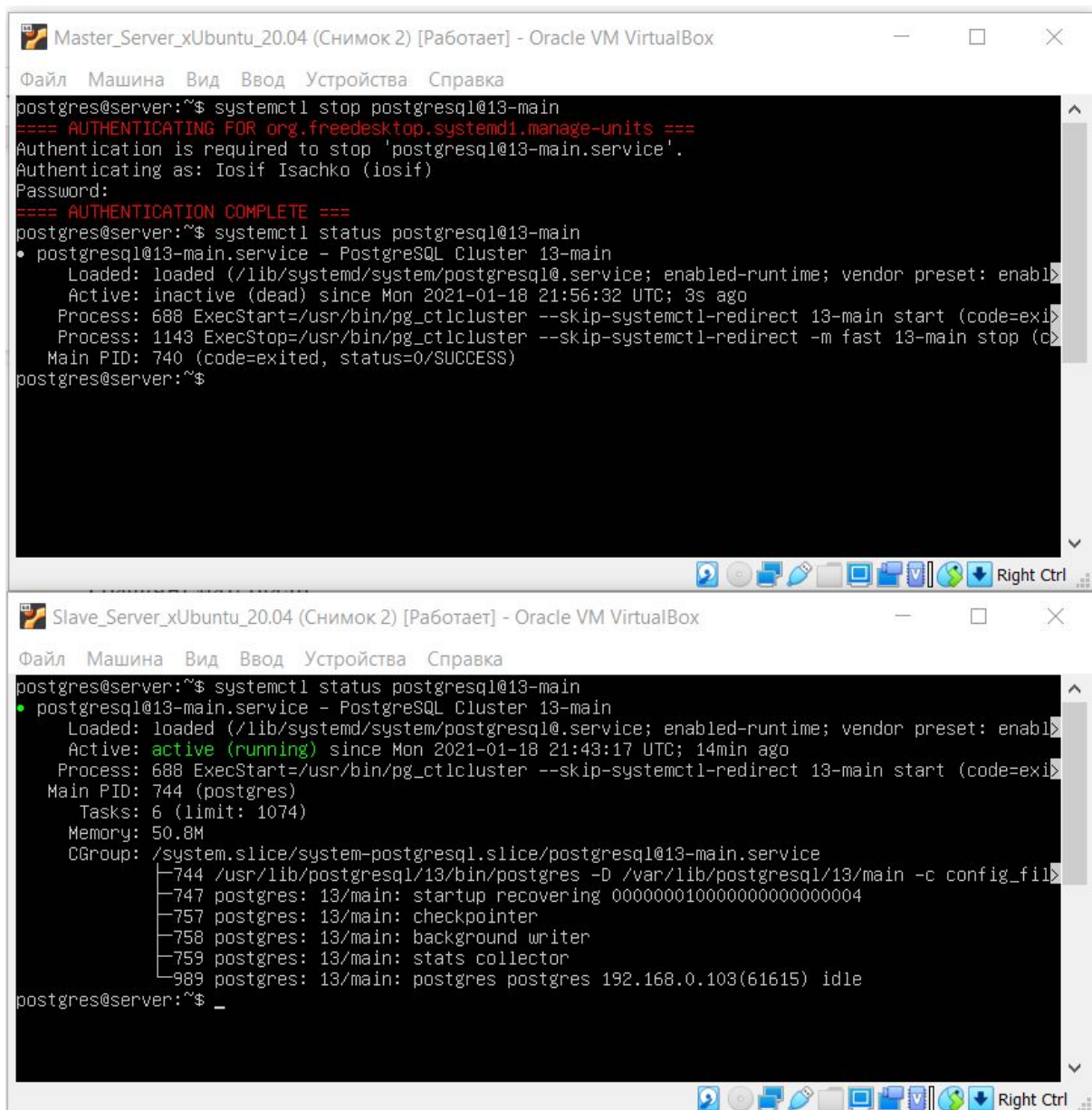


Рис. Ініціація випадку виникнення негараздів з головним сервером

command: search -c 528 -l 10
 Error query.
 Error: connect ECONNREFUSED 192.168.0.115:5432
 at TCPConnectWrap.afterConnect [as oncomplete] (net.js:1141:16)
 Error query.
 Error: connect ECONNREFUSED 192.168.0.115:5432
 at TCPConnectWrap.afterConnect [as oncomplete] (net.js:1141:16)
 Error query.
 Error: connect ECONNREFUSED 192.168.0.115:5432
 at TCPConnectWrap.afterConnect [as oncomplete] (net.js:1141:16)

(index)	id	title	price	date	category	goods_id
0	70887	'Компьютер Тест 2'	5000	2021-01-15T21:35:10.613Z	'Компьютеры(528)'	665467391
1	51239	'Компьютер HP 290 G1 MT (8PG33EA) Windows 10 Pro'	13999	2021-01-13T21:13:10.604Z	'Компьютеры(528)'	222121381
2	51277	'ARTLINE Business B21 v01 (B21v01)'	4218	2021-01-13T21:13:10.611Z	'Компьютеры(528)'	10099870
3	51265	'Компьютер Artline Business B12 v21 (B12v21)'	9480	2021-01-13T21:13:10.609Z	'Компьютеры(528)'	110464084
4	51242	'Компьютер HP 290 G1 MT (8PG31EA) Windows 10 Pro'	13599	2021-01-13T21:13:10.604Z	'Компьютеры(528)'	222120109
5	51236	'Монитор Lenovo ThinkCentre M820z (10SD05F00) Windows 10 Pro'	11999	2021-01-13T21:13:10.603Z	'Компьютеры(528)'	266138196
6	51267	'Монитор Acer Aspire C24-963 (DQ.BERME.006) Black/Silver'	24599	2021-01-13T21:13:10.609Z	'Компьютеры(528)'	223357345
7	51272	'Компьютер Artline Business B27 v34'	9396	2021-01-13T21:13:10.610Z	'Компьютеры(528)'	218694013
8	51287	'Компьютер Artline Business B29 v20'	12480	2021-01-13T21:13:10.613Z	'Компьютеры(528)'	154463408
9	51253	'Компьютер Artline Home H23 v01 (H23v01)'	8399	2021-01-13T21:13:10.606Z	'Компьютеры(528)'	42475336

Рис. Результат запиту у випадку, коли головний сервер перестав працювати