«Параллельные и распределённые вычисления» Лабораторная работа №8 Очередь задач	Выполнил:	Зенин М.А.
	Преподаватель:	Катаев А.В.
	Дата выполнения:	
	Подпись преподавателя	

Цель: Изучение работы с параллельными вычислениями в языке Python.

Задача: Реализовать совместную обработку данных группой исполнителей.

Скачивание статей гражданского кодекса с консультант плюс

Compose.yml с редисом

```yaml
services:
  redis:
    image: redis:6.2-alpine
    ports:
      - '6379:6379'
    volumes:
      - ./redis_data:/data
```

(запускается через docker compose up (-d))

Приложение celery

```python
from celery import Celery
import requests

app = Celery('tasks', broker='redis://localhost:6379/0')

@app.task
def download_link(url):
    headers = {
        'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Safari/537.36'
    }
    prefix = "https://www.consultant.ru"
    response = requests.get(prefix + url, headers=headers)

    if response.status_code == 200:
        with open(f"files/{url.replace("/", "-")}.html", "w") as file:
            file.write(response.text)
    else:
        print(url)
        print(f"Failed to retrieve the page. Status code: {response.status_code}")
```

Программка, отправляющая задачи:

```
from tasks import download_link
import requests
from bs4 import BeautifulSoup
from tqdm import tqdm

def extract_links(url):
    response = requests.get(url)

    if response.status_code == 200:
        soup = BeautifulSoup(response.text, 'html.parser')

        links = [a['href'] for a in soup.find_all('a', href=True) if not a['href'].startswith(('http://', 'https://', '//www'))]

        return links
    else:
        print(f"Failed to retrieve the page. Status code: {response.status_code}")
        return []

url = 'https://www.consultant.ru/document/cons_doc_LAW_5142/'
links = extract_links(url)

for link in tqdm(links):
    result = download_link.delay(link)
```

Запуск работника celery

```
(.venv) [zen@nex lab8]$ celery -A tasks worker --loglevel=info

 -------------- celery@nex v5.5.2 (immunity)
--- ***** -----
-- ******* ---- Linux-6.14.3-arch1-1-x86_64-with-glibc2.41 2025-04-30 05:18:31
- *** --- * ---
- ** ---------- [config]
- ** ---------- .> app:         tasks:0x7f3d0e021160
- ** ---------- .> transport:   redis://localhost:6379/0
- ** ---------- .> results:     disabled://
- *** --- * --- .> concurrency: 8 (prefork)
-- ******* ---- .> task events: OFF (enable -E to monitor tasks in this worker)
--- ***** -----
 -------------- [queues]
                .> celery           exchange=celery(direct) key=celery


[tasks]
  . tasks.download_link

[2025-04-30 05:18:31,731: INFO/MainProcess] Connected to redis://localhost:6379
/0
[2025-04-30 05:18:31,738: INFO/MainProcess] mingle: searching for neighbors
[2025-04-30 05:18:32,756: INFO/MainProcess] mingle: all alone
[2025-04-30 05:18:32,789: INFO/MainProcess] celery@nex ready.
```

Запущу несколько

```
-- ******* ---- Linux-6.14.3-arch1-1-x86_64-with-glibc2.41 2025-04-30 05:18:31
- *** --- * ---
- ** ---------- [config]
- ** ---------- .> app:         tasks:0x7f3d0e021160
- ** ---------- .> transport:   redis://localhost:6379/0
- ** ---------- .> results:     disabled://
- *** --- * --- .> concurrency: 8 (prefork)
-- ******* ---- .> task events: OFF (enable -E to monitor tasks in this worker)
--- ***** -----
-------------- [queues]
               .> celery           exchange=celery(direct) key=celery


[tasks]
  . tasks.download_link

[2025-04-30 05:18:31,731: INFO/MainProcess] Connected to redis://localhost:6379
/0
[2025-04-30 05:18:31,738: INFO/MainProcess] mingle: searching for neighbors
[2025-04-30 05:18:32,756: INFO/MainProcess] mingle: all alone
[2025-04-30 05:18:32,789: INFO/MainProcess] celery@nex ready.

- *** --- * ---
- ** ---------- [config]
- ** ---------- .> app:         tasks:0x795cd7bd9160
- ** ---------- .> transport:   redis://localhost:6379/0
- ** ---------- .> results:     disabled://
- *** --- * --- .> concurrency: 8 (prefork)
-- ******* ---- .> task events: OFF (enable -E to monitor tasks in this worker)
--- ***** -----
-------------- [queues]
               .> celery           exchange=celery(direct) key=celery


[tasks]
  . tasks.download_link

[2025-04-30 05:19:46,518: INFO/MainProcess] Connected to redis://localhost:6379
/0
[2025-04-30 05:19:46,525: INFO/MainProcess] mingle: searching for neighbors
[2025-04-30 05:19:47,546: INFO/MainProcess] mingle: all alone
[2025-04-30 05:19:47,576: INFO/MainProcess] celery@nex ready.
```

Выполняю программу, отправляющую задачи и работники начинают скачивать

Статьи:

```
-document-cons_doc_LAW_5142-f5742b6373038a15f928a5e6cdb18017bca5e62-.html
-document-cons_doc_LAW_5142-f589d9a4e0187211e24202a72b1d19b603fa3111-.html
-document-cons_doc_LAW_5142-f5e55adcc2ef06ee9e5b7f57f0f3ca3d2ea8825c-.html
-document-cons_doc_LAW_5142-f6393de25fd77addbf1c56c728db1da758416df8-.html
-document-cons_doc_LAW_5142-f670878d88ab83726bd1804b82668b84b027802e-.html
-document-cons_doc_LAW_5142-f6ea95d03afbe2db19cd1752b2e94173e4003cf4-.html
-document-cons_doc_LAW_5142-f6ef073fdeb7c69cfabc279056f1f4124956b5d9-.html
-document-cons_doc_LAW_5142-f71dd73496a05c5654fe07f42a96bb506f7ea5db-.html
-document-cons_doc_LAW_5142-f773bb7064ed1a1ceecfec079eb518cddabeb806-.html
-document-cons_doc_LAW_5142-f7871578ce9b026c450f64790704bd48c7d94bcb-.html
-document-cons_doc_LAW_5142-f7eff6cea3ee262cd97814d84d8780fe8e825462-.html
-document-cons_doc_LAW_5142-f7f26a277d8e29823e6b98ee86e2f33837a81450-.html
-document-cons_doc_LAW_5142-f8982a71849e23103bfd40e54eab7eae5bc87f9a-.html
-document-cons_doc_LAW_5142-f8f542de1a8540a32701afd8495edb3873597ece-.html
-document-cons_doc_LAW_5142-f9498350007fc6def03006086be05e29c6f32397-.html
-document-cons_doc_LAW_5142-f9658bf38181daad404c799a289d3df11ae5fd33-.html
-document-cons_doc_LAW_5142-f9732de88783800811973b3a13ef5112de0b5321-.html
-document-cons_doc_LAW_5142-f9d4f6974d23645aa246d991add4a4d39d601016-.html
-document-cons_doc_LAW_5142-fa1115a8e1086a097eb06c012bc2f5d5494b3756-.html
-document-cons_doc_LAW_5142-faa1d146348b1ced10eabfa204199bdd3ed1e98c-.html
-document-cons_doc_LAW_5142-fb394c3b45cbd2797e81fb4a1560dfb1225a664d-.html
-document-cons_doc_LAW_5142-fc48e160bcfdfeb0bba35922ae5b8377afc21404-.html
-document-cons_doc_LAW_5142-fc7edd948ce8489e4cf7532f2d485527525cbc73-.html
-document-cons_doc_LAW_5142-fd17bb3646b48877d28f1a445a17f12818031c9d-.html
-document-cons_doc_LAW_5142-fd6980fe6d3f891c65ca00040c9d100eeb2ef3ee-.html
-document-cons_doc_LAW_5142-fdd4b8c4accb6ef73badf6da6cd1124daa2a60e4-.html
```