

# Hướng dẫn cài đặt Redis và cập nhật dữ liệu lên Redis

## 1. Hướng dẫn cài đặt Redis trên Ubuntu

### Bước 1: Cập nhật hệ thống

Mở terminal và chạy lệnh sau để cập nhật danh sách gói phần mềm:

```
sudo apt update && sudo apt upgrade -y
```

### Bước 2: Cài đặt Redis

Ubuntu có sẵn Redis trong kho lưu trữ chính thức, bạn có thể cài đặt bằng lệnh:

```
sudo apt install redis-server -y
```

### Bước 3: Kiểm tra phiên bản Redis

Sau khi cài đặt, kiểm tra phiên bản Redis để đảm bảo quá trình cài đặt thành công:

```
redis-server --version
```

### Bước 4: Cấu hình Redis

Mở file cấu hình Redis để chỉnh sửa:

```
sudo nano /etc/redis/redis.conf
```

Tìm dòng:

```
supervised no
```

Sửa thành

```
supervised system
```

Điều này giúp Redis chạy dưới dạng dịch vụ systemd.

Ngoài ra, nếu bạn muốn Redis lắng nghe trên tất cả các IP (ví dụ để kết nối từ máy khác), tìm:

```
bind 127.0.0.1 ::1
```

Và thay bằng:

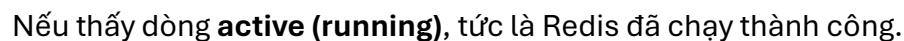
```
bind 0.0.0.0
```

Sau khi hoàn tất, nhấn CTRL + X, chọn Y và nhấn Enter để lưu.

### Bước 5: Khởi động lại Redis

```
sudo systemctl restart redis
```

```
sudo systemctl status redis
```



```
sudo systemctl enable redis
```

Các bước thực hiện:

## Bước 1: Cài đặt các thư viện cần thiết

Trước khi chạy script, cần đảm bảo Redis đã được cài đặt trên Ubuntu (xem hướng dẫn cài đặt ở trên) và bạn cần cài đặt thư viện **Redis-Py** và **Pandas**:

```
pip install redis pandas
```

Ngoài ra bạn cũng có thể sử dụng thêm thư viện gc nhằm các mục đích như giảm tiêu thụ bộ nhớ, tăng hiệu suất và ngăn rò rỉ bộ nhớ.

## Bước 2: Chuẩn bị dữ liệu

Script sử dụng ba file CSV để upload dữ liệu lên Redis:

- **anime-dataset-2023.csv**: Chứa thông tin về các bộ anime.
- **users-details-2023.csv**: Chứa thông tin cá nhân người dùng.
- **users-score-2023.csv**: Chứa thông tin điểm đánh giá anime của người dùng.

## Bước 3: Cấu hình kết nối Redis

Mở file **ReadData.py** và kiểm tra thông tin kết nối Redis:

```
redis_host = "192.168.126.131" # Địa chỉ IP của Redis server  
redis_port = 6379 # Cổng mặc định của Redis
```

Nếu Redis chạy trên **localhost**, hãy sửa `redis_host = "127.0.0.1"`.

## Bước 4: Chạy script để upload dữ liệu

Chạy file **ReadData.py** bằng Python để tiến hành Upload dữ liệu:

### Lưu thông tin Anime vào Redis

- Đọc file anime-dataset-2023.csv theo từng phần (chunk\_size=50000) để tiết kiệm bộ nhớ.
- Dữ liệu được lưu dưới dạng **Redis Hash**, mỗi anime có một key kiểu anime:{anime\_id}.

```
r.hset(f"anime:{anime_id}", mapping=row)
```

### Lưu thông tin cá nhân người dùng vào Redis

- Đọc file users-details-2023.csv theo từng phần (chunk\_size=50000).
- Mỗi người dùng có một key kiểu user:{user\_id} được lưu dưới dạng **Redis Hash**.

```
r.hset(f"user:{user_id}", mapping=row)
```

### Lưu điểm đánh giá người dùng

- Đọc file users-score-2023.csv theo từng phần (chunk\_size=100000).
- Lưu điểm đánh giá theo cặp user\_id\_anime\_id dưới dạng **Redis String**.

```
r.set(f"{user_id}_{anime_id}", rating)
```

### Bước 5: Kiểm tra dữ liệu trong Redis

Sau khi chạy script, bạn có thể kiểm tra Redis bằng Redis CLI.

#### Kiểm tra số lượng Key hiện có trong Database

```
redis-cli
```

```
dbsize
```

```
127.0.0.1:6379> dbsize
(integer) 25081386
```

#### Kiểm tra thông tin một anime

```
redis-cli
```

```
hgetall anime:1 # Kiểm tra thông tin anime có id = 1
```

```
11) "Aired"
12) "Apr 3, 1998 to Apr 24, 1999"
13) "Synopsis"
14) "Crime is timeless. By the year 2071, humanity has expanded across the galaxy, filling the surface of other planets with settlements like those on Earth. These new societies are plagued by murder, drug use, and theft, and intergalactic outlaws are hunted by a growing number of tough bounty hunters.\n\nSpike Spiegel and Jet Black pursue criminals throughout space to make a humble living. Beneath his goofy and aloof demeanor, Spike is haunted by the weight of his violent past. Meanwhile, Jet manages his own troubled memories while taking care of Spike and the Bebop, their ship. The duo is joined by the beautiful con artist Faye Valentine, odd child Edward Wong Hau Pepelu Tivrusky IV, and Ein, a bioengineered Welsh Corgi.\n\nWhile developing bonds and working to catch a colorful cast of criminals, the Bebop crew's lives are disrupted by a menace from Spike's past. As a rival's maniacal plot continues to unravel, Spike must choose between life with his newfound family or revenge for his old wounds."
15) "Studios"
16) "Sunrise"
17) "Members"
18) "1771505"
19) "Rank"
20) "41.0"
21) "Rating"
22) "R - 17+ (violence & profanity)"
23) "Episodes"
24) "26.0"
25) "Licensors"
26) "Funimation, Bandai Entertainment"
27) "anime_id"
28) "1"
29) "Popularity"
```

## Kiểm tra điểm đánh giá của một user

get 1\_49 # Kiểm tra điểm đánh giá của user 1 cho anime có anime\_id là 49

```
127.0.0.1:6379> GET 1_49
"8"
```

## Kiểm tra thông tin một user

hgetall user:1 # Kiểm tra thông tin user có id = 1

```
127.0.0.1:6379> hgetall user:1
1) "Mal ID"
2) "1"
3) "Username"
4) "Xinil"
5) "Gender"
6) "Male"
7) "Birthday"
8) "1985-03-04T00:00:00+00:00"
9) "Location"
10) "California"
11) "Joined"
12) "2004-11-05T00:00:00+00:00"
13) "Days Watched"
14) "142.3"
15) "Mean Score"
16) "7.37"
17) "Watching"
18) "1.0"
19) "Completed"
20) "233.0"
21) "On Hold"
22) "8.0"
23) "Dropped"
24) "93.0"
25) "Plan to Watch"
26) "64.0"
27) "Total Entries"
```