Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
УРАЛЬСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ ПЕРВОГО ПРЕЗИДЕНТА РОССИИ Б.Н. ЕЛЬЦИНА
(УрФУ имени первого Президента России Б.Н. Ельцина)
Институт радиоэлектроники и информационных технологий — РТФ

ОТЧЁТ

по лабораторной работе №1

по дисциплине «Методы и инструменты анализа больших данных»

Преподаватель _____ (дата) _____ (подпись) С.Г. Мирвода

Студент _____ (дата) _____ (подпись) А.М. Белоусов

Студент _____ (дата) _____ (подпись) А.В. Жиденко

Группа: РИМ-201211

Екатеринбург 2021

Задание 1

1. Воспользовавшись примерами из презентации (Лекция 5 - Основы хранилищ ключ-значение) установите redis

$ sudo apt-get update

```
root@ubuntu:/home/belozhi# sudo apt-get update
Get:1 http://security.ubuntu.com/ubuntu focal-security InRelease [114 kB]
Hit:2 http://us.archive.ubuntu.com/ubuntu focal InRelease
Get:3 http://us.archive.ubuntu.com/ubuntu focal-updates InRelease [114 kB]
Get:4 http://us.archive.ubuntu.com/ubuntu focal-backports InRelease [108 kB]
Fetched 336 kB in 1s (228 kB/s)
Reading package lists... Done
```

$ sudo apt-get install redis-server

```
root@ubuntu:/home/belozhi# sudo apt-get install redis-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libatomic1 libhiredis0.14 libjemalloc2 liblua5.1-0 lua-bitop lua-cjson
  redis-tools
Suggested packages:
  ruby-redis
The following NEW packages will be installed:
  libatomic1 libhiredis0.14 libjemalloc2 liblua5.1-0 lua-bitop lua-cjson
  redis-server redis-tools
0 upgraded, 8 newly installed, 0 to remove and 144 not upgraded.
Need to get 925 kB of archives.
After this operation, 4 123 kB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us.archive.ubuntu.com/ubuntu focal-updates/main amd64 libatomic1 amd64 10.3.0-1ubuntu1~20.04 [9 284 B]
Get:2 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 libhiredis0.14 amd64 0.14.0-6 [30,2 kB]
Get:3 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 libjemalloc2 amd64 5.2.1-1ubuntu1 [235 kB]
Get:4 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 liblua5.1-0 amd64 5.1.5-8.1build4 [99,9 kB]
Get:5 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 lua-bitop amd64 1.0.2-5 [6 680 B]
Get:6 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 lua-cjson amd64 2.1.0+dfsg-2.1 [17,4 kB]
Get:7 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-tools amd64 5:5.0.7-2 [489 kB]
Get:8 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server amd64 5:5.0.7-2 [37,3 kB]
Fetched 925 kB in 1s (691 kB/s)
Selecting previously unselected package libatomic1:amd64.
(Reading database ... 205798 files and directories currently installed.)
```

$ redis-server

```
root@ubuntu:/home/belozhi# redis-server
4980:C 04 Dec 2021 13:25:19.236 # oO0OoO0OoO0Oo Redis is starting oO0OoO0OoO0Oo
4980:C 04 Dec 2021 13:25:19.236 # Redis version=5.0.7, bits=64, commit=00000000, modified=0, pid=4980, just started
4980:C 04 Dec 2021 13:25:19.236 # Warning: no config file specified, using the default config. In order to specify a config file use redis-server /pat
h/to/redis.conf
4980:M 04 Dec 2021 13:25:19.236 * Increased maximum number of open files to 10032 (it was originally set to 1024).
4980:M 04 Dec 2021 13:25:19.237 # Could not create server TCP listening socket *:6379: bind: Address already in use
```

$ service redis-server status

```
root@ubuntu:/home/belozhi# service redis-server status
● redis-server.service - Advanced key-value store
     Loaded: loaded (/lib/systemd/system/redis-server.service; enabled; vendor preset: enabled)
     Active: active (running) since Sat 2021-12-04 13:20:20 +05; 7min ago
       Docs: http://redis.io/documentation,
             man:redis-server(1)
   Main PID: 4197 (redis-server)
      Tasks: 4 (limit: 19092)
     Memory: 1.8M
     CGroup: /system.slice/redis-server.service
             └─4197 /usr/bin/redis-server 127.0.0.1:6379

дек 04 13:20:20 ubuntu systemd[1]: Starting Advanced key-value store...
дек 04 13:20:20 ubuntu systemd[1]: redis-server.service: Can't open PID file /run/redis/redis-server.pid (yet?) after start: Operation not permitted
дек 04 13:20:20 ubuntu systemd[1]: Started Advanced key-value store.
```

$ systemctl status redis

```
root@ubuntu:/home/belozhi# systemctl status redis
● redis-server.service - Advanced key-value store
     Loaded: loaded (/lib/systemd/system/redis-server.service; enabled; vendor preset: enabled)
     Active: active (running) since Sat 2021-12-04 13:20:20 +05; 8min ago
       Docs: http://redis.io/documentation,
             man:redis-server(1)
   Main PID: 4197 (redis-server)
      Tasks: 4 (limit: 19092)
     Memory: 1.8M
     CGroup: /system.slice/redis-server.service
             └─4197 /usr/bin/redis-server 127.0.0.1:6379

дек 04 13:20:20 ubuntu systemd[1]: Starting Advanced key-value store...
дек 04 13:20:20 ubuntu systemd[1]: redis-server.service: Can't open PID file /run/redis/redis-server.pid (yet?) after start: Operation not permitted
дек 04 13:20:20 ubuntu systemd[1]: Started Advanced key-value store.
```

2. Попробуйте все команды, которые приведены в презентации, для справки необходимо использовать документацию по командам https://redis.io/commands.

$ redis-cli

Подключение

```
root@ubuntu:/home/belozhi# redis-cli
127.0.0.1:6379> PING
PONG
127.0.0.1:6379> PING xxx
"xxx"
```

Конфигурация

```
127.0.0.1:6379> config get loglevel
1) "loglevel"
2) "notice"
127.0.0.1:6379> config get *
  1) "dbfilename"
  2) "dump.rdb"
  3) "requirepass"
  4) ""
  5) "masterauth"
  6) ""
  7) "cluster-announce-ip"
  8) ""
  9) "unixsocket"
 10) ""
 11) "logfile"
 12) "/var/log/redis/redis-server.log"
 13) "pidfile"
 14) "/var/run/redis/redis-server.pid"
 15) "slave-announce-ip"
 16) ""
 17) "replica-announce-ip"
 18) ""
```

```
127.0.0.1:6379> config set loglevel "notice"
OK
```

Простые типы данных

```
127.0.0.1:6379> set users:msg "{name: Alexander, likes: [redis]}"
OK
127.0.0.1:6379> strlen users:msg
(integer) 33
127.0.0.1:6379> getrange users:msg 23 27
": [re"
127.0.0.1:6379> append users:msg " FTW !!111"
(integer) 43
127.0.0.1:6379> get users:msg
"{name: Alexander, likes: [redis]} FTW !!111"
```

```
127.0.0.1:6379> set users:msg "{name: AndreBelo, likes: [redis]}"
OK
127.0.0.1:6379> strlen users:msg
(integer) 33
127.0.0.1:6379> getrange users:msg 23 27
": [re"
127.0.0.1:6379> append users:msg " AlexLexaDLB"
(integer) 45
127.0.0.1:6379> get users:msg
"{name: AndreBelo, likes: [redis]} AlexLexaDLB"
```

Terminal - root@ubuntu: /home/belozhi

File Edit View Terminal Tabs Help

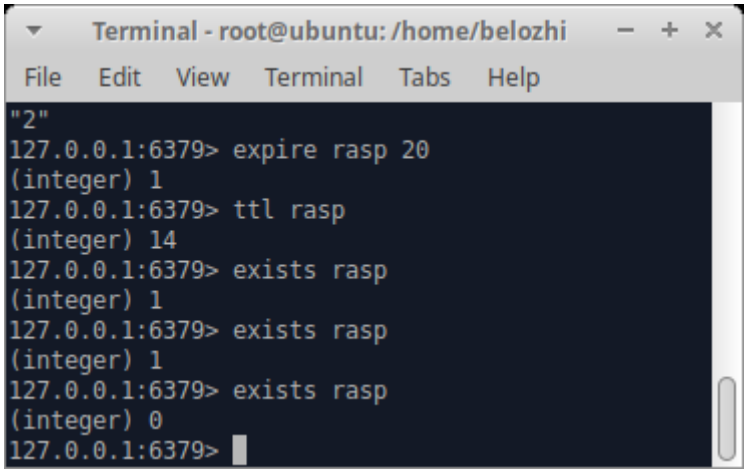
```
(integer) 45
127.0.0.1:6379> get users:msg
"{name: AndreBelo, likes: [redis]} AlexLexaDLB"
127.0.0.1:6379> set rasp http://urfu.ru/ru/students/study/schedule/
OK
127.0.0.1:6379> get rasp
"http://urfu.ru/ru/students/study/schedule/"
127.0.0.1:6379>
```

Terminal - root@ubuntu: /home/belozhi

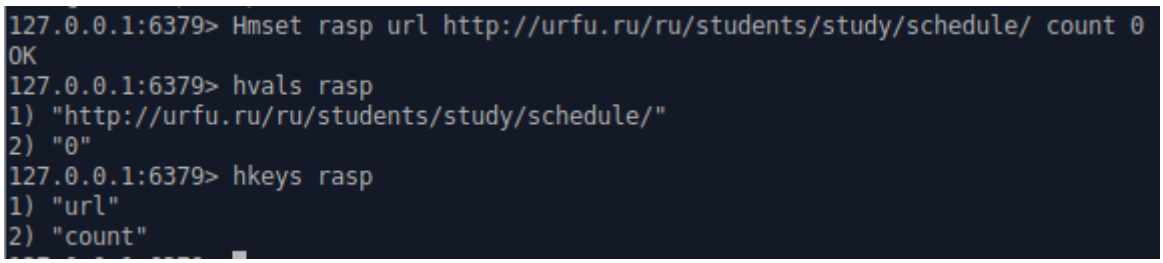
File Edit View Terminal Tabs Help

```
"http://urfu.ru/ru/students/study/schedule/"
127.0.0.1:6379> set rasp:count 3
OK
127.0.0.1:6379> set rasp:count 0
OK
127.0.0.1:6379> incr rasp:count
(integer) 1
127.0.0.1:6379> incr rasp:count
(integer) 2
127.0.0.1:6379> get rasp:count
"2"
127.0.0.1:6379>
```

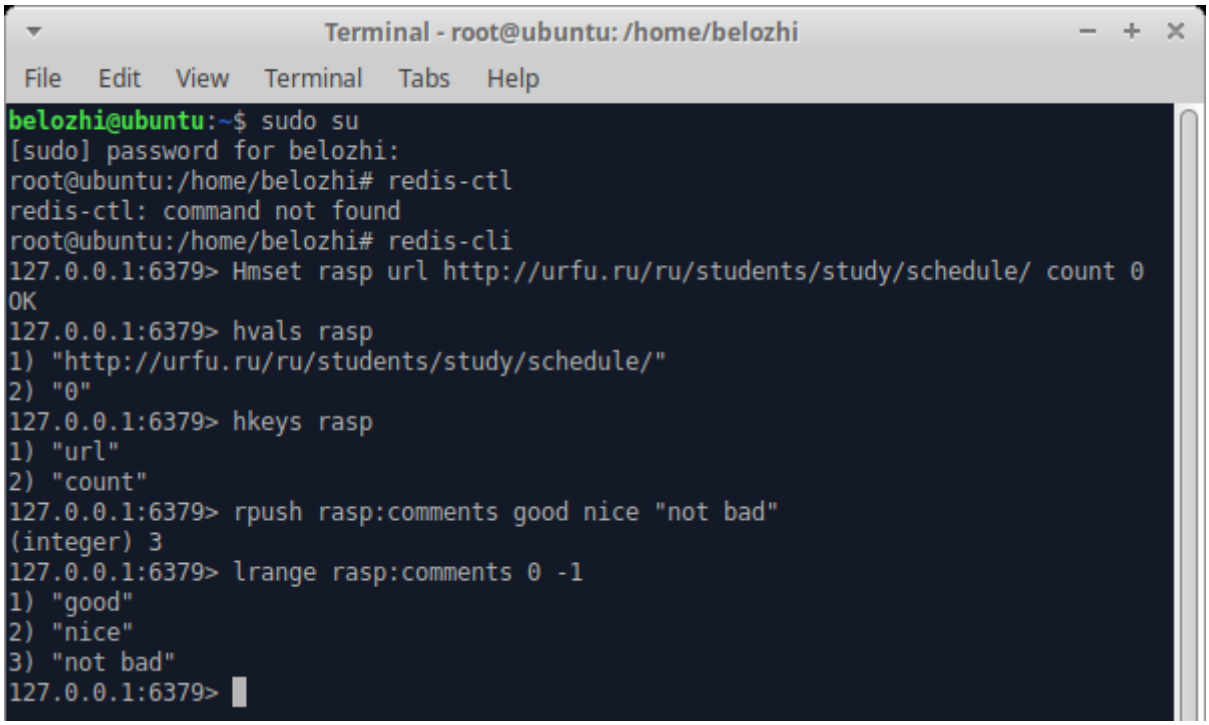
Устаревание



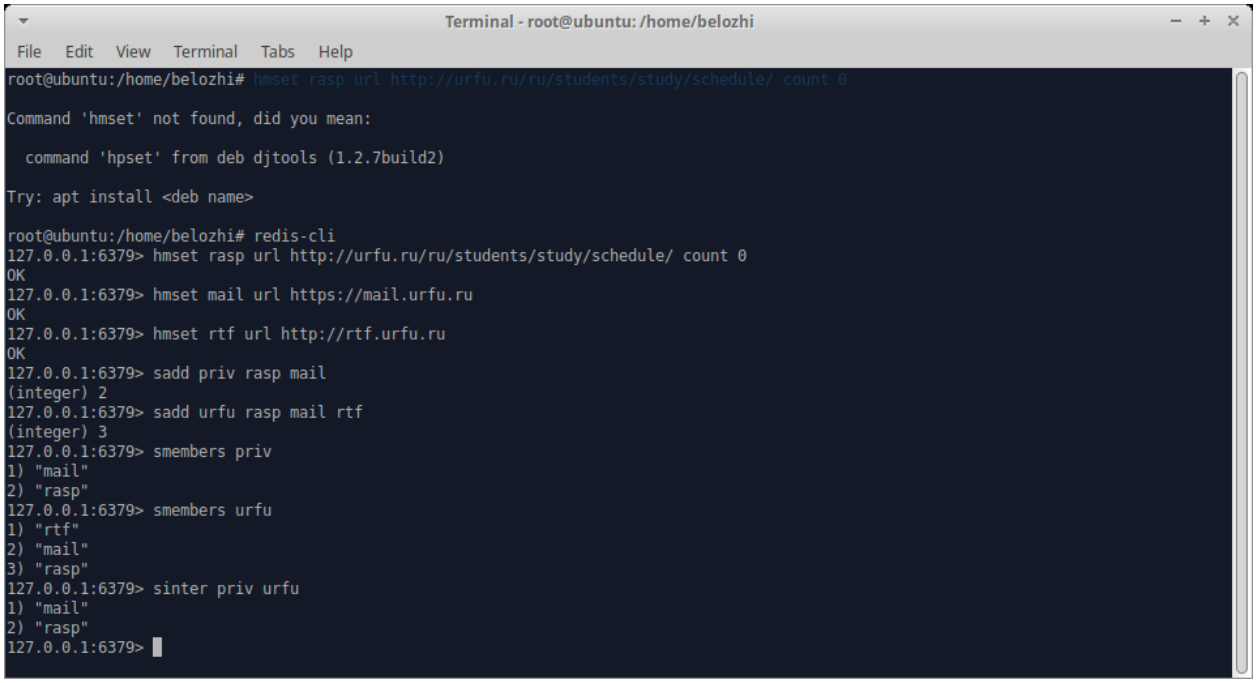
Хэши



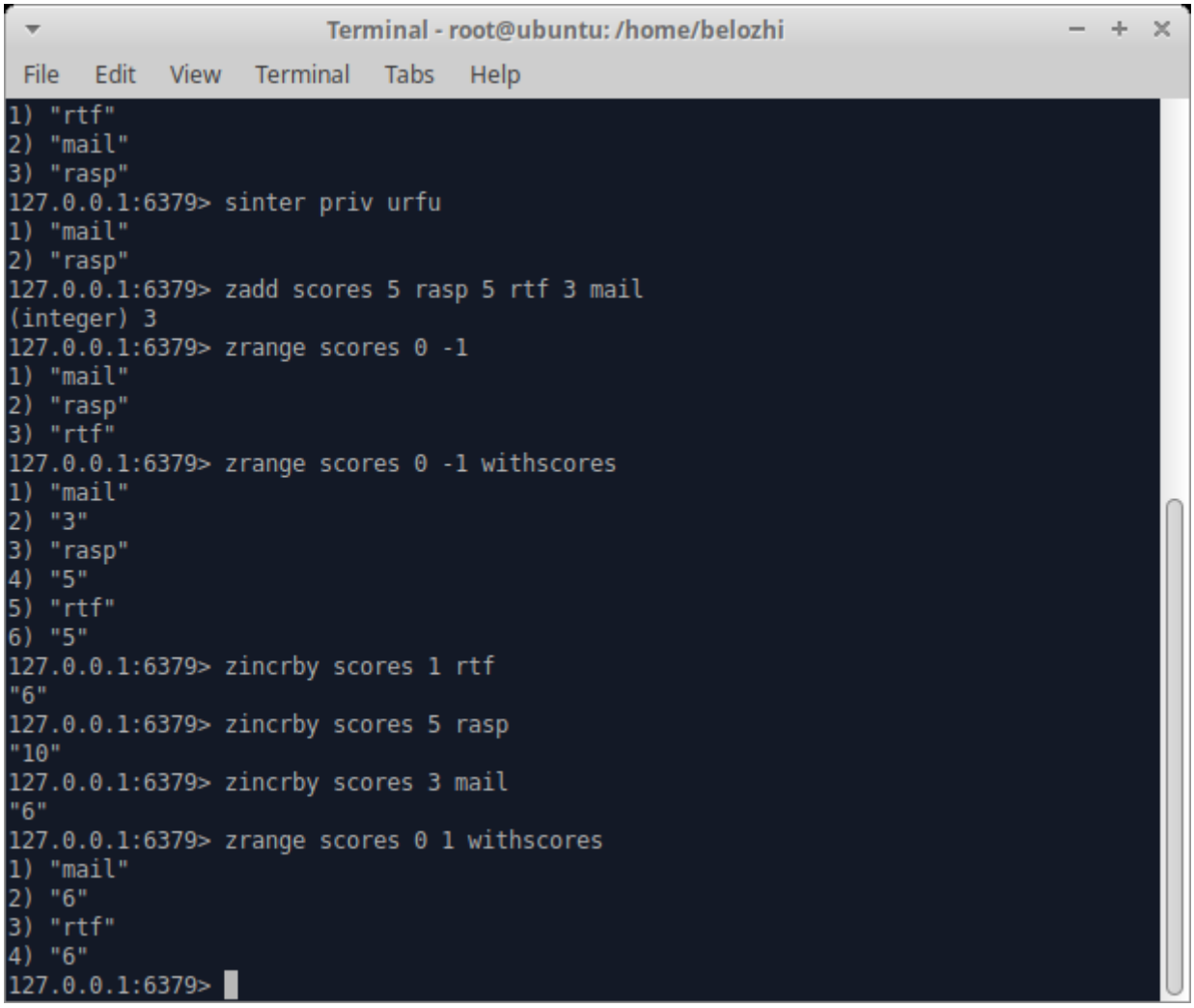
Списки



Множества и объединения



Упорядоченные множества



Задание 2

1. После выполнения первого задания выполните скрипт

```
belozhi@ubuntu:~$ #!/bin/bash
belozhi@ubuntu:~$
belozhi@ubuntu:~$ REDIS_KEY_PATTERN="${REDIS_KEY_PATTERN:-*}"
belozhi@ubuntu:~$ for key in $(redis-cli --scan --pattern "$REDIS_KEY_PATTERN")
> do
>     type=$(redis-cli type $key)
>     if [ $type = "list" ]
>     then
>         printf "$key => \n$(redis-cli lrange $key 0 -1 | sed 's/^/  /')\n"
>     elif [ $type = "hash" ]
>     then
>         printf "$key => \n$(redis-cli hgetall $key | sed 's/^/  /')\n"
>     else
>         printf "$key => $(redis-cli get $key)\n"
>     fi
> done
rasp:count => 2
urfu => WRONGTYPE Operation against a key holding the wrong kind of value
scores => WRONGTYPE Operation against a key holding the wrong kind of value
rtf =>
  url
  http://rtf.urfu.ru
users:msg => {name: AndreBelo, likes: [redis]} AlexLexaDLB
mail =>
  url
  https://mail.urfu.ru
rasp:comments =>
  good
  nice
  not bad
rasp =>
  url
  http://urfu.ru/ru/students/study/schedule/
  count
  0
priv => WRONGTYPE Operation against a key holding the wrong kind of value
```

Сохраните вывод этого скрипта в отельный файл, это будет результатом первого задания.

```
belozhi@ubuntu:~$ for key in $(redis-cli --scan --pattern "$REDIS_KEY_PATTERN"); do     type=$(re
dis-cli type $key);     if [ $type = "list" ];     then         printf "$key => \n$(redis-cli lra
nge $key 0 -1 | sed 's/^/  /')\n";     elif [ $type = "hash" ];     then         printf "$key =>
\n$(redis-cli hgetall $key | sed 's/^/  /')\n";     else         printf "$key => $(redis-cli get
$key)\n";     fi; done >belozhi.txt
belozhi@ubuntu:~$ ls -l belozhi.txt
-rw-rw-r-- 1 belozhi belozhi 486 дек  4 14:14 belozhi.txt
belozhi@ubuntu:~$ cat belozhi.txt
rasp:count => 2
urfu => WRONGTYPE Operation against a key holding the wrong kind of value
scores => WRONGTYPE Operation against a key holding the wrong kind of value
rtf =>
  url
  http://rtf.urfu.ru
users:msg => {name: AndreBelo, likes: [redis]} AlexLexaDLB
mail =>
  url
  https://mail.urfu.ru
rasp:comments =>
  good
  nice
  not bad
rasp =>
  url
  http://urfu.ru/ru/students/study/schedule/
  count
  0
priv => WRONGTYPE Operation against a key holding the wrong kind of value
belozhi@ubuntu:~$
```

Результат в файле belozhi.txt

```
Terminal - belozhi@ubuntu: ~
File  Edit  View  Terminal  Tabs  Help
  GNU nano 4.8                          belozhi.txt
rasp:count => 2
urfu => WRONGTYPE Operation against a key holding the wrong kind of value
scores => WRONGTYPE Operation against a key holding the wrong kind of value
rtf =>
  url
  http://rtf.urfu.ru
users:msg => {name: AndreBelo, likes: [redis]} AlexLexaDLB
mail =>
  url
  https://mail.urfu.ru
rasp:comments =>
  good
  nice
  not bad
rasp =>
  url
  http://urfu.ru/ru/students/study/schedule/
  count
  0
priv => WRONGTYPE Operation against a key holding the wrong kind of value


^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify     ^C Cur Pos
^X Exit        ^R Read File   ^\ Replace     ^U Paste Text  ^T To Spell    ^_ Go To Line
```

Как видно из текста, скрипт использует команды scan, type, lrange, hgetall, get для получения с помощью консольного клиента redis-cli списка ключей, получения типа ключа, и получения значения ключей в зависимости от типа (списки, хэши, простые ключи).

2. Напишите скрипт для установки 1M простых ключей, 10M простых ключей замерьте скорость и потребление памяти

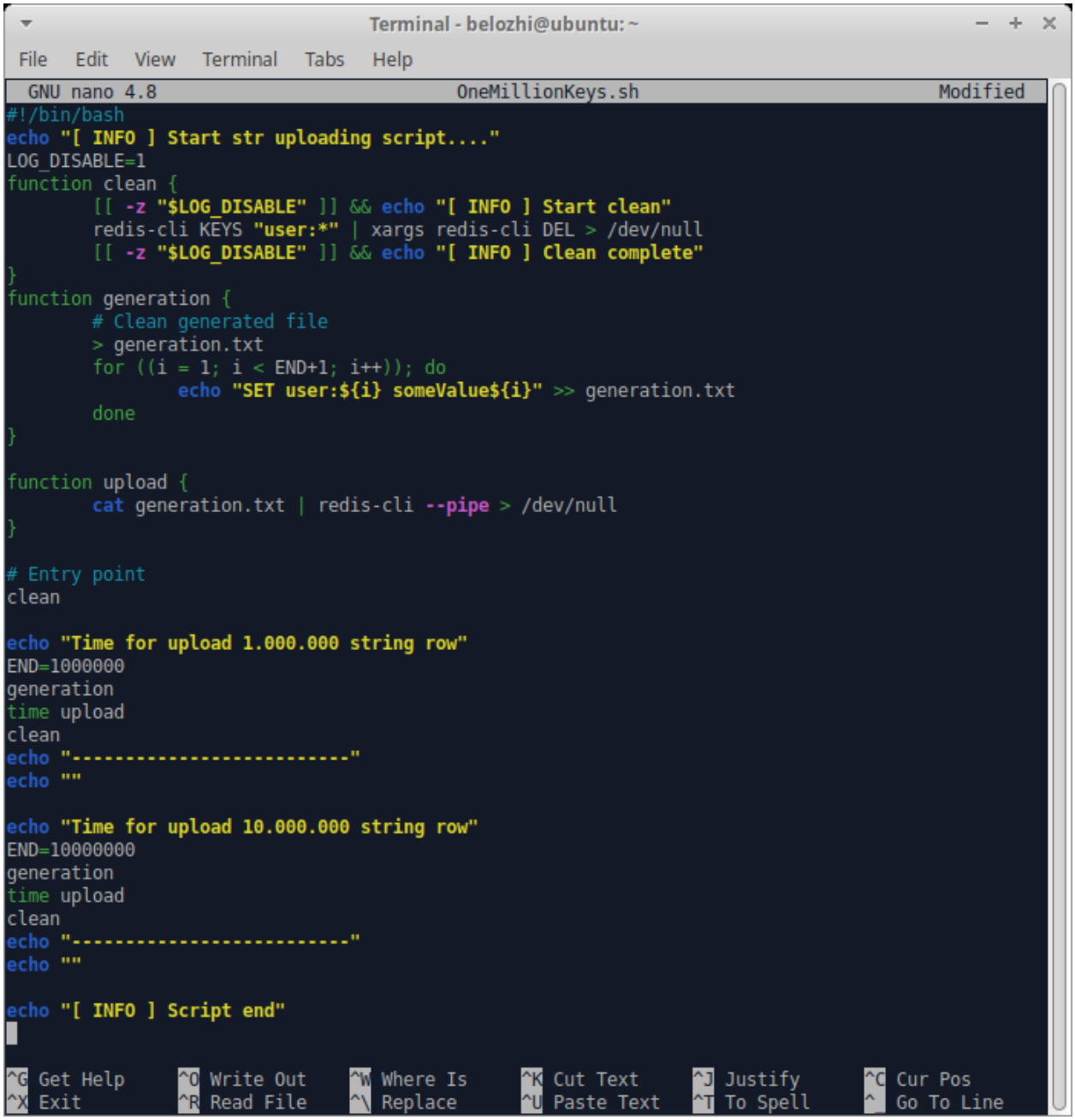


Таблица 1 – Замеры времени и ОЗУ при записи ключей

Count	Time №1	Time №2	Time №3	RAM
1.000.000	0m2,842s	0m3,034s	0m3,088s	136 Mb
10.000.000	0m38,120s	0m36,675s	0m37,425s	938 Mb

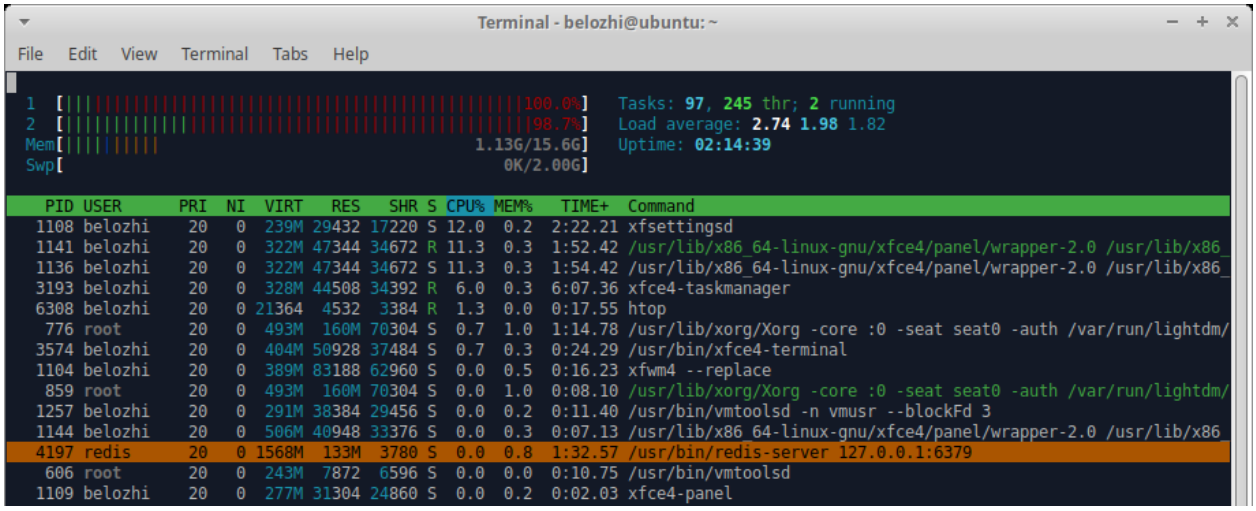


Рисунок 1 – Загрузка ОЗУ при записи 10M ключей

3. Напишите скрипт для установки 1M хэшей, 10M хэшей. Замерьте скорость и потребление памяти

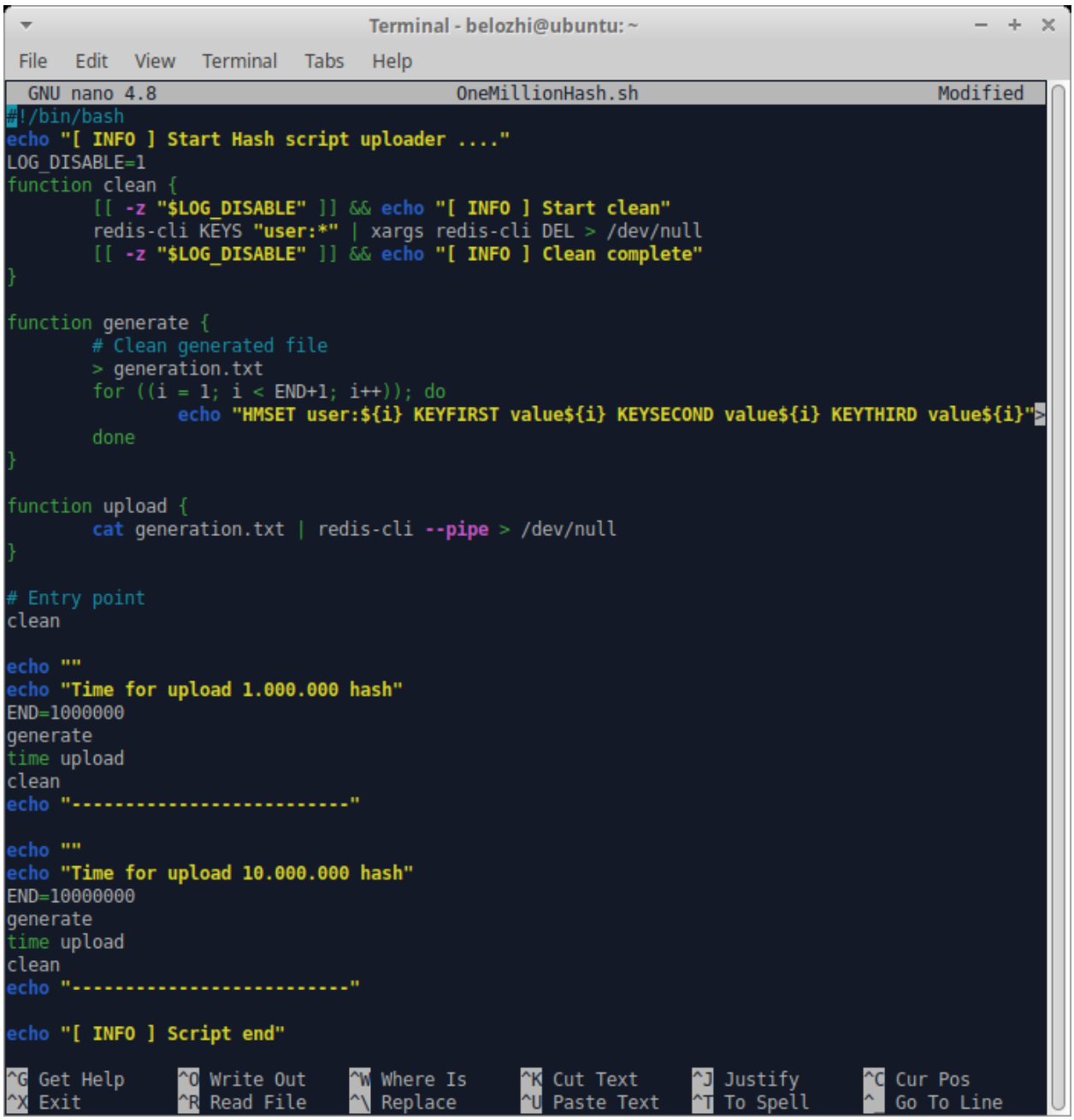


Таблица 2 – Замеры времени и ОЗУ при записи хэшей

Count	Time №1	Time №2	Time №3	RAM
1.000.000	0m6,666s	0m6,457s	0m6,380s	187 Mb
10.000.000	1m21,008s	1m24,012s	1m20,831s	1627 Mb

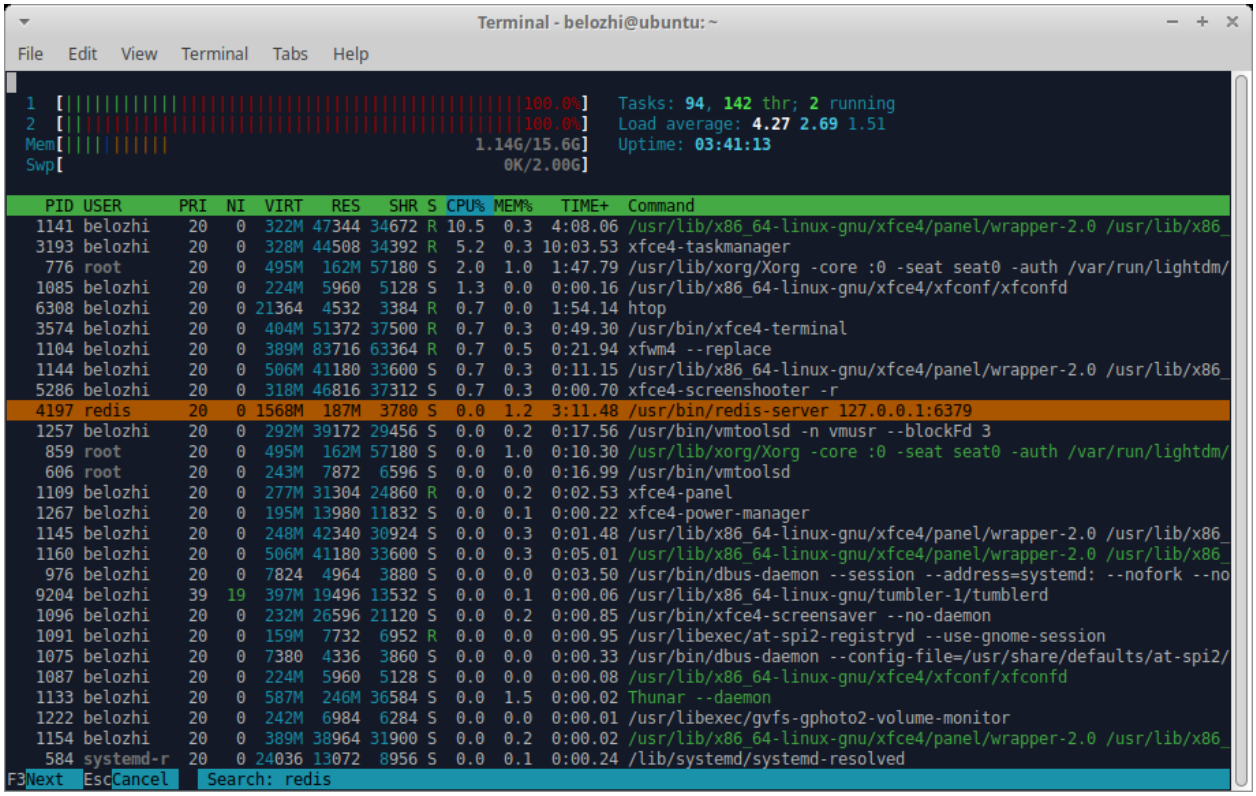
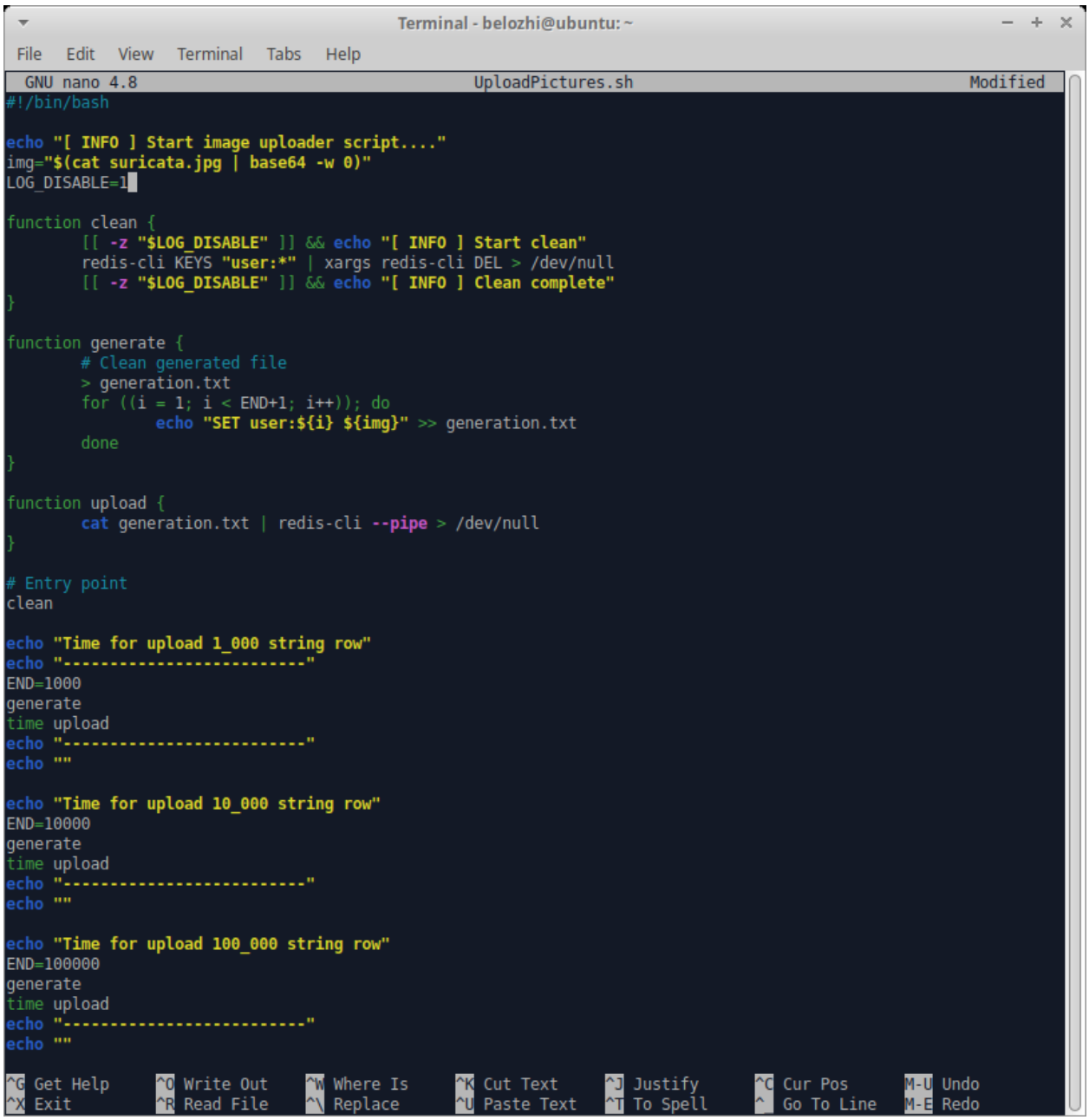


Рисунок 2 – Загрузка ОЗУ при записи 10M хэшей

4. Сохраните на диск аватарку из любой социальной сети и, воспользовавшись результатами из предыдущих пунктов, рассчитайте сколько таких картинок может вместить ОЗУ вашего компьютера, модифицируйте скрипт и попробуйте записать в redis в виде простых ключей. Посмотрите результаты.

Рисунок 3 – Фото для загрузки в Redis



Count	Time	RAM
1000	0m0,671s	32 Mb
10 000	0m6,368s	325 Mb
100 000	1m16,636s	4343 Mb
1 000 000	6m49,268s	11,1 Gb

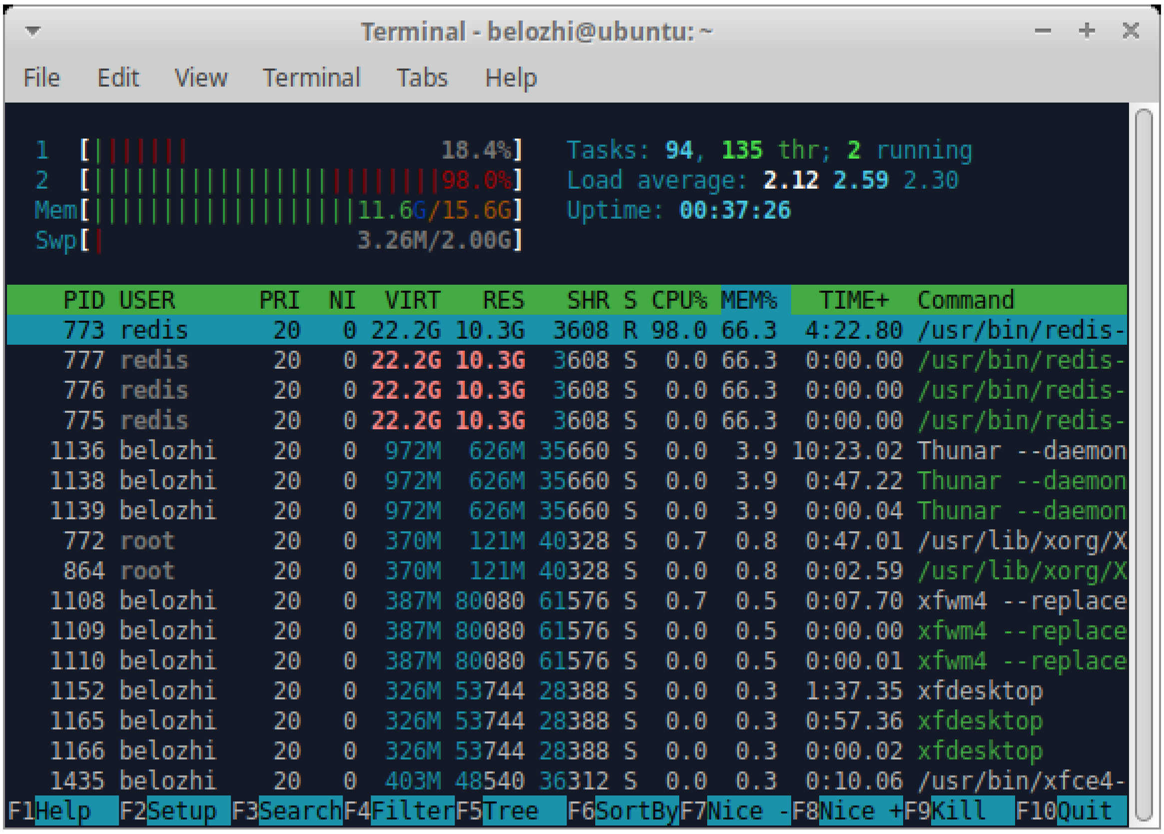


Рисунок 4 – Загрузка ОЗУ при записи 1М картинок

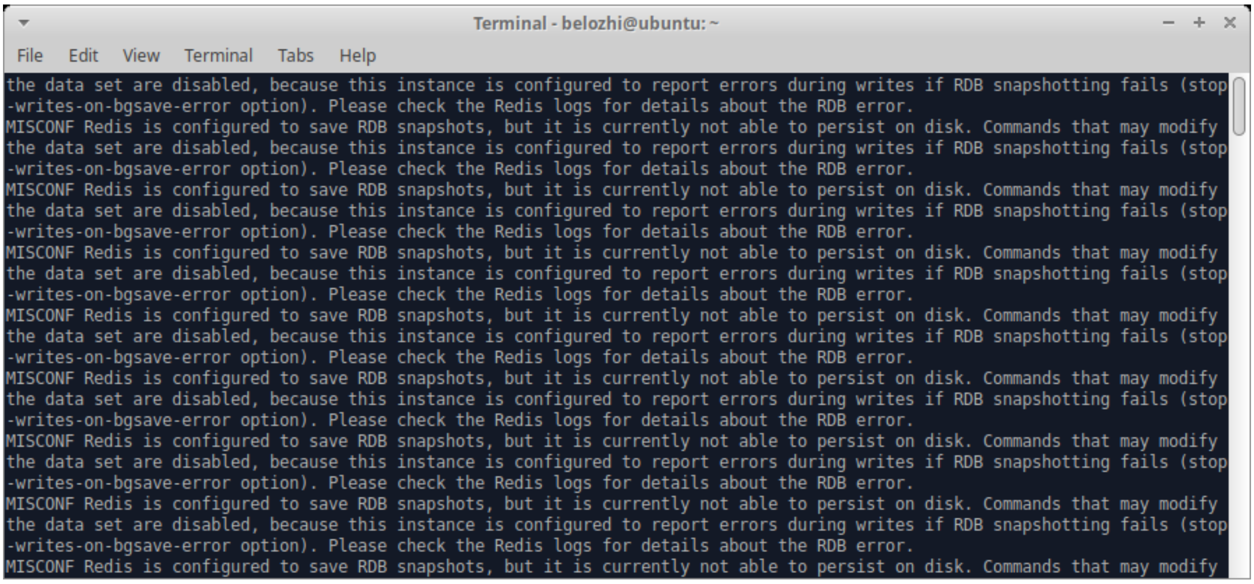


Рисунок 5 – Ошибка при записи 1М картинок