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			С.Г. Мирвода
	(дата)	(подпись)	ситипрыда
Студент			А.М. Белоусов
	(дата)	(подпись)	J
Студент			А.В. Жиденко
•	(дата)	(подпись)	

Группа: РИМ-201211

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-lubuntu1~20.04 [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 liblua5.1-0 amd64 5.1.5-8.lbuild4 [99,9 kB]
Get:4 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 lua-bitop amd64 5.1.5-8.lbuild4 [99,9 kB]
Get:5 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 lua-cjson amd64 2.1.0-dfsg-2.1 [17,4 kB]
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 lua-cjson amd64 5:5.0.7-2 [489 kB]
Get:6 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server amd64 5:5.0.7-2 [489 kB]
Get:7 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server amd64 5:5.0.7-2 [47,4 kB]
Get:8 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server amd64 5:5.0.7-2 [47,4 kB]
Get:8 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server 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 [487,4 kB]
Get:8 http://us.archive.ubuntu.com/ubuntu focal/universe amd64 redis-server amd64 5:5.0.7-2 [487,4 kB]
Get:8 http://us.ar
```

\$ redis-server

```
root@ubuntu:/home/belozhi# redis-server
4980:C 04 Dec 2021 13:25:19.236 # o00000000000 Redis is starting o00000000000
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

\$ 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
L4197 /usr/bin/redis-server 127.0.0.1:6379

qek 04 13:20:20 ubuntu systemd[1]: Starting Advanced key-value store...
qek 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

    "loglevel"

2) "notice"
127.0.0.1:6379> config get *
 1) "dbfilename"
 2) "dump.rdb"
 3) "requirepass"
 5) "masterauth"
  7) "cluster-announce-ip"
 9) "unixsocket"
 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"
```

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]}"
0K
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]}"
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"
```

```
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/

0K

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> 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

```
Terminal - root@ubuntu: /home/belozhi
 File
      Edit View Terminal Tabs
                                   Help
'2"
127.0.0.1:6379> expire rasp 20
(integer) 1
127.0.0.1:6379> ttl rasp
(integer) 14
127.0.0.1:6379> exists rasp
(integer) 1
127.0.0.1:6379> exists rasp
(integer) 1
127.0.0.1:6379> exists rasp
(integer) 0
127.0.0.1:6379>
```

Хэши

```
127.0.0.1:6379> Hmset rasp url http://urfu.ru/ru/students/study/schedule/ count 0
OK
127.0.0.1:6379> hvals rasp
1) "http://urfu.ru/ru/students/study/schedule/"
2) "0"
127.0.0.1:6379> hkeys rasp
1) "url"
2) "count"
```

Списки

```
Terminal - root@ubuntu: /home/belozhi
                                                                                 - +
 File
      Edit
            View
                  Terminal Tabs
                                    Help
belozhi@ubuntu:~$ sudo su
[sudo] password for belozhi:
root@ubuntu:/home/belozhi# redis-ctl
redis-ctl: command not found
root@ubuntu:/home/belozhi# redis-cli
127.0.0.1:6379> Hmset rasp url http://urfu.ru/ru/students/study/schedule/ count 0
127.0.0.1:6379> hvals rasp
1) "http://urfu.ru/ru/students/study/schedule/"
2) "0"
127.0.0.1:6379> hkeys rasp
1) "url"
2) "count"
127.0.0.1:6379> rpush rasp:comments good nice "not bad"
(integer) 3
127.0.0.1:6379> lrange rasp:comments 0 -1
1) "good"
2) "nice"
3) "not bad"
127.0.0.1:6379>
```

Множества и объединения

```
Terminal-root@ubuntu:/home/belozhi# mset rasp url http://urfu.ru/ru/students/study/schedule/ count 0

Command 'hmset' not found, did you mean:
    command 'hmset' from deb djtools (1.2.7build2)

Try: apt install <deb name>

root@ubuntu:/home/belozhi# redis-cli
127.0.0.1:6379> hmset rasp url http://urfu.ru/ru/students/study/schedule/ count 0

OK
127.0.0.1:6379> hmset mail url https://mail.urfu.ru

OK
127.0.0.1:6379> hmset rtf url http://rtf.urfu.ru

OK
127.0.0.1:6379> sadd priv rasp mail
(integer) 2
127.0.0.1:6379> smembers priv
(integer) 3
127.0.0.1:6379> smembers priv
1) "masil"
2) "rasp"
127.0.0.1:6379> smembers urfu
1) "rasp"
127.0.0.1:6379> smembers urfu
1) "rasp"
127.0.0.1:6379> sinter priv urfu
1) "mail"
2) "masil"
2) "masil"
2) "rasp"
127.0.0.1:6379> sinter priv urfu
1) "mail"
2) "rasp"
127.0.0.1:6379> sinter priv urfu
1) "mail"
2) "rasp"
127.0.0.1:6379> sinter priv urfu
1) "mail"
2) "rasp"
127.0.0.1:6379> sinter priv urfu
```

Упорядоченные множества

```
Terminal - root@ubuntu: /home/belozhi
                                                                                     + ×
 File
      Edit
           View
                   Terminal Tabs Help

 "rtf"

2) "mail"
3) "rasp"
127.0.0.1:6379> sinter priv urfu
1) "mail"
2) "rasp"
127.0.0.1:6379> zadd scores 5 rasp 5 rtf 3 mail
(integer) 3
127.0.0.1:6379> zrange scores 0 -1
1) "mail"
2) "rasp"
3) "rtf"
127.0.0.1:6379> zrange scores 0 -1 withscores
1) "mail"
2) "3"
3) "rasp"
4) "5"
5) "rtf"
6) "5"
127.0.0.1:6379> zincrby scores 1 rtf
"6"
127.0.0.1:6379> zincrby scores 5 rasp
"10"
127.0.0.1:6379> zincrby scores 3 mail
"6"
127.0.0.1:6379> zrange scores 0 1 withscores
1) "mail"
2) "6"
3) "rtf"
4) "6"
127.0.0.1:6379>
```

Задание 2

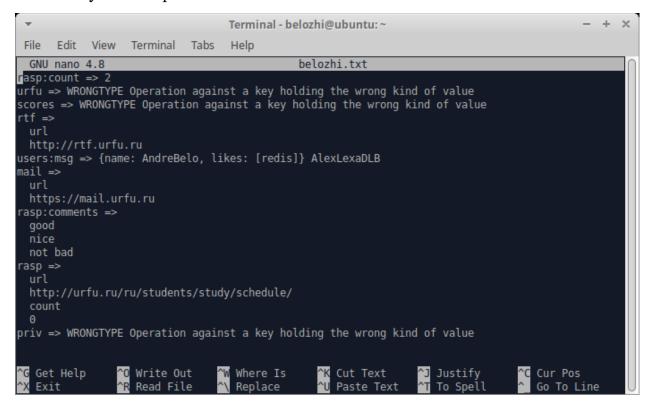
1. После выполнения первого задания выполните скрипт

```
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")
      type=$(redis-cli type $key)
      if [ $type = "list" ]
      printf "$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 =>
  http://urfu.ru/ru/students/study/schedule/
 count
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
nge $key 0 -1 | sed 's/^/ /')\n";    elif [ $type = "hash" ];
\n$(redis-cli hgetall $key | sed 's/^/ /')\n";    else
                                                                      printf "$key => \n$(redis-cli lra
                                                                                         printf "$key =>
                                                                        then
                                                                        printf "$key => $(redis-cli get
               fi; done >belozhi.txt
$key)\n";
belozhi@ubuntu:~$ 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
  not bad
rasp =>
  url
  http://urfu.ru/ru/students/study/schedule/
priv => WRONGTYPE Operation against a key holding the wrong kind of value
belozhi@ubuntu:~$
```

Результат в файле belozhi.txt



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

```
Terminal - belozhi@ubuntu: ~
                                                                                                            - + ×
 File Edit View Terminal Tabs
GNU nano 4.8
                                                  OneMillionKeys.sh
                                                                                                        Modified
echo "[ INFO ] Start str uploading script...."
LOG DISABLE=1
         [[ -z "$LOG_DISABLE" ]] && echo "[ INFO ] Start clean"
redis-cli KEYS "user:*" | xargs redis-cli DEL > /dev/null
[[ -z "$LOG_DISABLE" ]] && echo "[ INFO ] Clean complete"
function generation {
         > generation.txt
         for ((i = 1; i < END+1; i++)); do
echo "SET user:\{i\} someValue\{i\}" >>> generation.txt
function upload {
         cat generation.txt | redis-cli --pipe > /dev/null
clean
echo "Time for upload 1.000.000 string row"
END=1000000
generation
time upload
echo "Time for upload 10.000.000 string row"
END=10000000
generation
time upload
clean
echo "----"
echo ""
echo "[ INFO ] Script end"
                                                                             ^J Justify
                                      ^W Where Is
                                                         ^K Cut Text
                   ^O Write Out
^B Read File
                                                                                                ^C Cur Pos
^G Get Help
                                          Replace
```

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

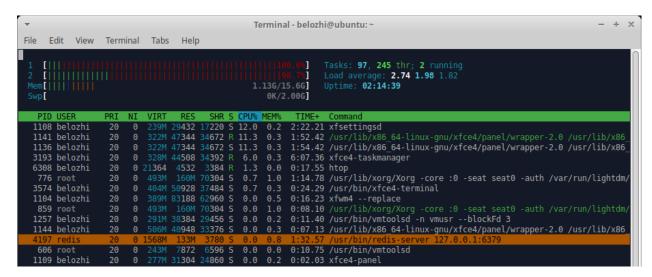


Рисунок 1 – Загрузка ОЗУ при записи 10М ключей

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

```
- + ×
                                      Terminal - belozhi@ubuntu: ~
 File Edit View Terminal Tabs Help
                                               OneMillionHash.sh
                                                                                                  Modified
 GNU nano 4.8
cho "[ INFO ] Start Hash script uploader ...."
LOG DISABLE=1
function clean {
        [[ -z "$LOG_DISABLE" ]] && echo "[ INFO ] Start clean"
redis-cli KEYS "user:*" | xargs redis-cli DEL > /dev/null
[[ -z "$LOG_DISABLE" ]] && echo "[ INFO ] Clean complete"
function generate {
        > generation.txt
        for ((i = 1; i < END+1; i++)); do
                echo "HMSET user:${i} KEYFIRST value${i} KEYSECOND value${i} KEYTHIRD value${i}"
function upload {
         cat generation.txt | redis-cli --pipe > /dev/null
clean
echo "Time for upload 1.000.000 hash"
END=1000000
generate
time upload
clean
echo "Time for upload 10.000.000 hash"
END=10000000
generate
time upload
clean
echo "-----"
echo "[ INFO ] Script end"
                                                                        ^J Justify
^T To Spell
^G Get Help
                  ^O Write Out
                                    ^W Where Is
                                                         Cut Text
                                                                                           ^C Cur Pos
                     Read File
                                       Replace
                                                         Paste Text
                                                                                              Go To Line
   Exit
```

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

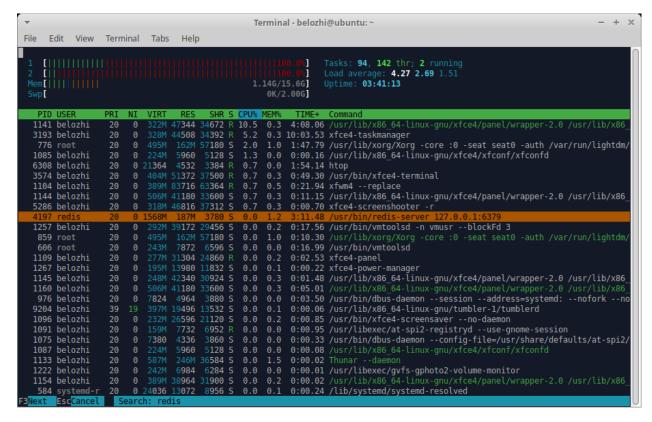
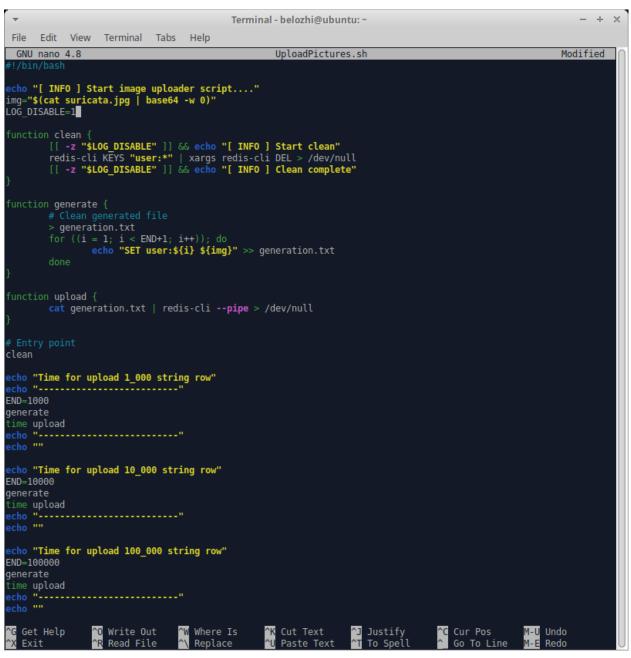


Рисунок 2 – Загрузка ОЗУ при записи 10М хэшей

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

```
Terminal - belozhi@ubuntu: ~
                                                                                   30
File
     Edit
           View
                  Terminal
                            Tabs
                                   Help
                               18.4%
                                         Tasks: 94, 135 thr; 2 running
                                         Load average: 2.12 2.59 2.30
                           .6G/15.6G1
                                         Uptime: 00:37:26
                         3.26M/2.00G
  PID USER
                          VIRT
                                 RES
                 PRI
                      NI
                                        SHR S CPU% MEM%
                                                            TIME+
                                                                   Command
                          22.2G 10.3G
                                                           4:22.80 /usr/bin/redis
   773 redis
                                        3608 R 98.0 66.3
                                                           0:00.00 /usr/bin/redis
                        0 22.2G 10.3G
   777 redis
                   20
                                        3608 S
                                                0.0 66.3
   776 redis
                   20
                        0 22.2G 10.3G
                                        3608 S
                                                0.0 66.3
                                                           0:00.00 /usr/bin/redis
   775 redis
                   20
                        0 22.2G 10.3G
                                        3608 S
                                                0.0 66.3
                                                          0:00.00
                                                                   /usr/bin/redis
  1136 belozhi
                                 626M 35660 S
                   20
                                                0.0
                                                     3.9 10:23.02 Thunar --daemon
                        0
                                                          0:47.22 Thunar --daemor
0:00.04 Thunar --daemor
  1138 belozhi
                   20
                                 626M 35660 S
                                                     3.9
                        0
                                                0.0
  1139 belozhi
                   20
                                 626M 35660 S
                                                     3.9
                        0
                                                0.0
                   20
                                 121M 40328 S
                                                          0:47.01 /usr/lib/xorg/X
   772 root
                        0
                                                0.7
                                                     0.8
                   20
                                 121M 40328 S
                                                          0:02.59 /usr/lib/xorg/
  864 root
                        0
                           370M
                                                0.0
                                                     0.8
  1108 belozhi
                           387M 80080 61576 S
                                                     0.5
                                                          0:07.70 xfwm4 --replace
                   20
                        0
                                                0.7
                                                          0:00.00 xfwm4 --replace
  1109 belozhi
                           387M 80080 61576 S
                                                     0.5
                   20
                        0
                                                0.0
                                                          0:00.01 xfwm4 --replace
  1110 belozhi
                           387M 80080 61576 S
                                                     0.5
                   20
                        0
                                                0.0
  1152 belozhi
                           326M 53744 28388 S
                                                           1:37.35 xfdesktop
                   20
                        0
                                                0.0
                                                     0.3
  1165 belozhi
                           326M 53744 28388
                   20
                        0
                                                0.0
                                                     0.3
                                                          0:57.36 xfdesktop
                           326M 53744 28388
  1166 belozhi
                   20
                        0
                                                0.0
                                                     0.3
                                                          0:00.02 xfdesktop
                           403M 48540 36312 S
  1435 belozhi
                   20
                        0
                                                0.0
                                                     0.3
                                                          0:10.06 /usr/bin/xfce4
      F2Setup F3SearchF4FilterF5Tree F6SortByF7Nice -F8Nice +F9Kill
```

Рисунок 4 – Загрузка ОЗУ при записи 1М картинок

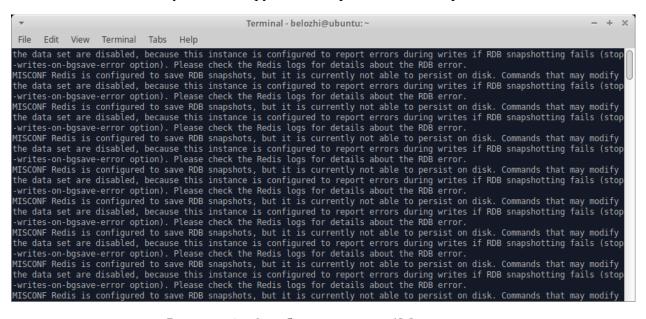


Рисунок 5 – Ошибка при записи 1М картинок