Linux tutorial: basic apps and utilities

Nikita Neveditsin, SMU, 2019

nikita.neveditsin@smu.ca

File Transfer

Most popular file transfer protocols used in LINUX: **SCP** and **SFTP**

Both are secure and use underlying SSH protocol (Secure Shell (SSH) is a cryptographic network protocol for operating network services securely over an *unsecured* network. SSH provides a secure channel over an unsecured network in a client-server architecture, connecting an SSH client application with an SSH server. [1]). Version 2 of SSH is now used. Default port: 22

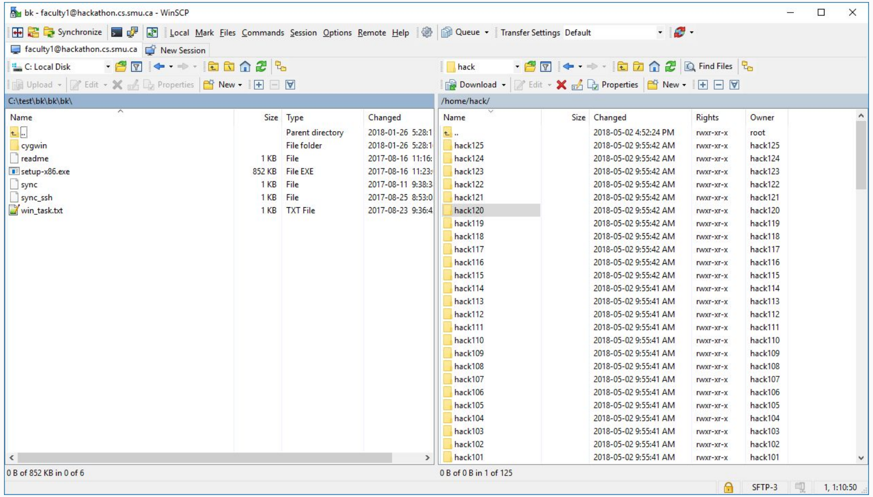
The SCP is a network protocol, which supports file transfers between hosts on a network. SCP uses SSH for data transfer and uses the same mechanisms for authentication, thereby ensuring the authenticity and confidentiality of the data in transit. A client can send (upload) files to a server, optionally including their basic attributes (permissions, timestamps). Clients can also request files or directories from a server (download). SCP runs over TCP port 22 by default [2]

Compared to the SCP protocol, which only allows file transfers, the SFTP (SSH File Transfer Protocol) allows for a range of operations on remote files which make it more like a remote file system protocol. An SFTP client's extra capabilities include resuming interrupted transfers, directory listings, and remote file removal.[3] Based on SSH, also uses port 22 by default.

SFTP != FTP over SSH! It’s a different protocol! Use SFTP where you can

File Transfer

WinSCP: most popular SCP/SFTP client for Windows



bk - faculty1@hackathon.cs.smu.ca - WinSCP

Transfer Settings: Default

faculty1@hackathon.cs.smu.ca | New Session

C: Local Disk

C:\test\bk\bk\bk\

Name	Size	Type	Changed
..		Parent directory	2018-01-26 5:28:1
cygwin		File folder	2018-01-26 5:28:1
readme	1 KB	File	2017-08-16 11:16:
setup-x86.exe	852 KB	File EXE	2017-08-16 11:23:
sync	1 KB	File	2017-08-11 9:38:3
sync_ssh	1 KB	File	2017-08-25 8:53:0
win_task.txt	1 KB	TXT File	2017-08-23 9:36:4

0 B of 852 KB in 0 of 6

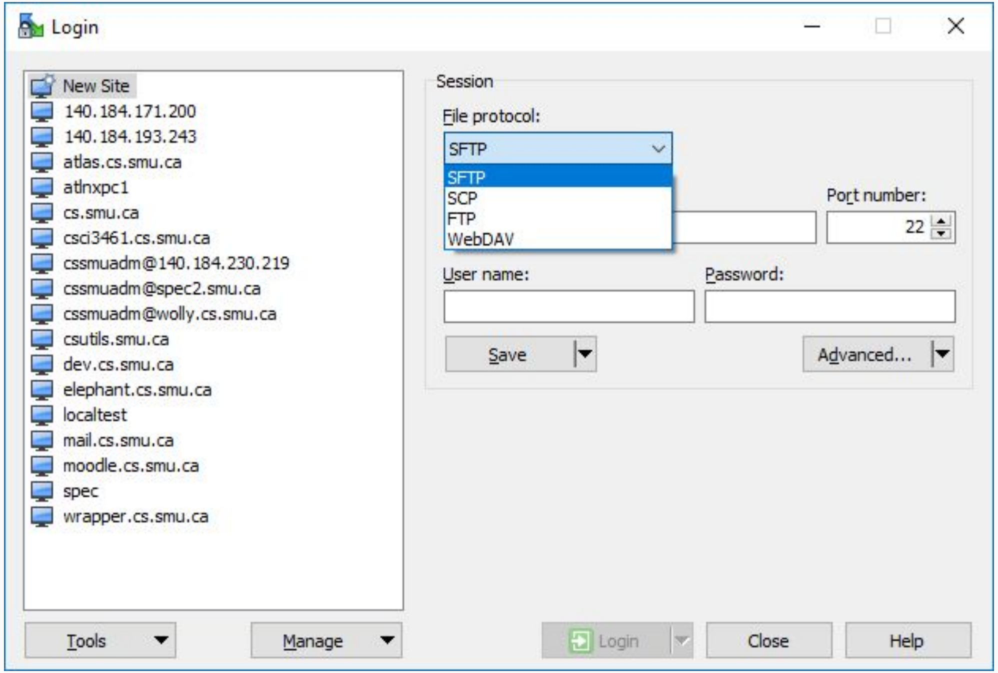
hack

/home/hack/

Name	Size	Changed	Rights	Owner
..		2018-05-02 4:52:24 PM	rwxr-xr-x	root
hack125		2018-05-02 9:55:42 AM	rwxr-xr-x	hack125
hack124		2018-05-02 9:55:42 AM	rwxr-xr-x	hack124
hack123		2018-05-02 9:55:42 AM	rwxr-xr-x	hack123
hack122		2018-05-02 9:55:42 AM	rwxr-xr-x	hack122
hack121		2018-05-02 9:55:42 AM	rwxr-xr-x	hack121
hack120		2018-05-02 9:55:42 AM	rwxr-xr-x	hack120
hack119		2018-05-02 9:55:42 AM	rwxr-xr-x	hack119
hack118		2018-05-02 9:55:42 AM	rwxr-xr-x	hack118
hack117		2018-05-02 9:55:42 AM	rwxr-xr-x	hack117
hack116		2018-05-02 9:55:42 AM	rwxr-xr-x	hack116
hack115		2018-05-02 9:55:42 AM	rwxr-xr-x	hack115
hack114		2018-05-02 9:55:41 AM	rwxr-xr-x	hack114
hack113		2018-05-02 9:55:41 AM	rwxr-xr-x	hack113
hack112		2018-05-02 9:55:41 AM	rwxr-xr-x	hack112
hack111		2018-05-02 9:55:41 AM	rwxr-xr-x	hack111
hack110		2018-05-02 9:55:41 AM	rwxr-xr-x	hack110
hack109		2018-05-02 9:55:41 AM	rwxr-xr-x	hack109
hack108		2018-05-02 9:55:41 AM	rwxr-xr-x	hack108
hack107		2018-05-02 9:55:41 AM	rwxr-xr-x	hack107
hack106		2018-05-02 9:55:41 AM	rwxr-xr-x	hack106
hack105		2018-05-02 9:55:41 AM	rwxr-xr-x	hack105
hack104		2018-05-02 9:55:41 AM	rwxr-xr-x	hack104
hack103		2018-05-02 9:55:41 AM	rwxr-xr-x	hack103
hack102		2018-05-02 9:55:41 AM	rwxr-xr-x	hack102
hack101		2018-05-02 9:55:41 AM	rwxr-xr-x	hack101

0 B of 0 B in 1 of 125

SFTP-3 | 1, 1:10:50



Login

- New Site
- 140.184.171.200
- 140.184.193.243
- atlas.cs.smu.ca
- atlnxpc1
- cs.smu.ca
- csci3461.cs.smu.ca
- cssmuadm@140.184.230.219
- cssmuadm@spec2.smu.ca
- cssmuadm@wolly.cs.smu.ca
- csutils.smu.ca
- dev.cs.smu.ca
- elephant.cs.smu.ca
- localtest
- mail.cs.smu.ca
- moodle.cs.smu.ca
- spec
- wrapper.cs.smu.ca

Session

File protocol: SFTP

- SFTP
- SCP
- FTP
- WebDAV

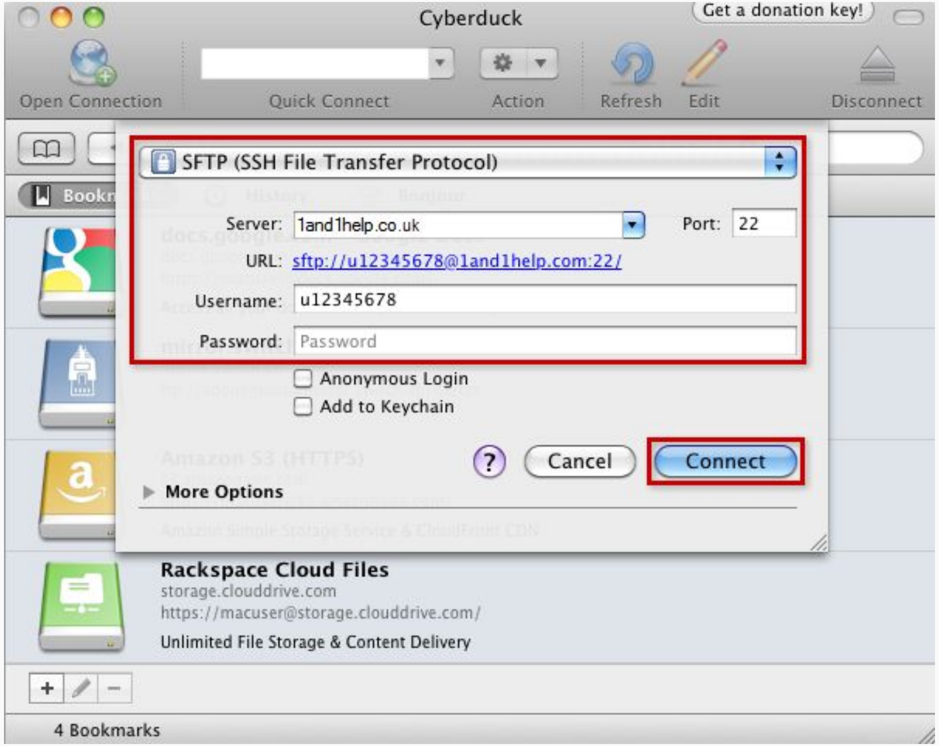
Port number: 22

User name:

Password:

File Transfer

For MAC OS users: Cyberduck



SFTP (SSH File Transfer Protocol)

Server: 1and1help.co.uk Port: 22

URL: sftp://u12345678@1and1help.com:22/

Username: u12345678

Password: Password

□ Anonymous Login

□ Add to Keychain

Rackspace Cloud Files
storage.clouddrive.com
https://macuser@storage.clouddrive.com/
Unlimited File Storage & Content Delivery

4 Bookmarks

[PL1]

File Transfer

scp

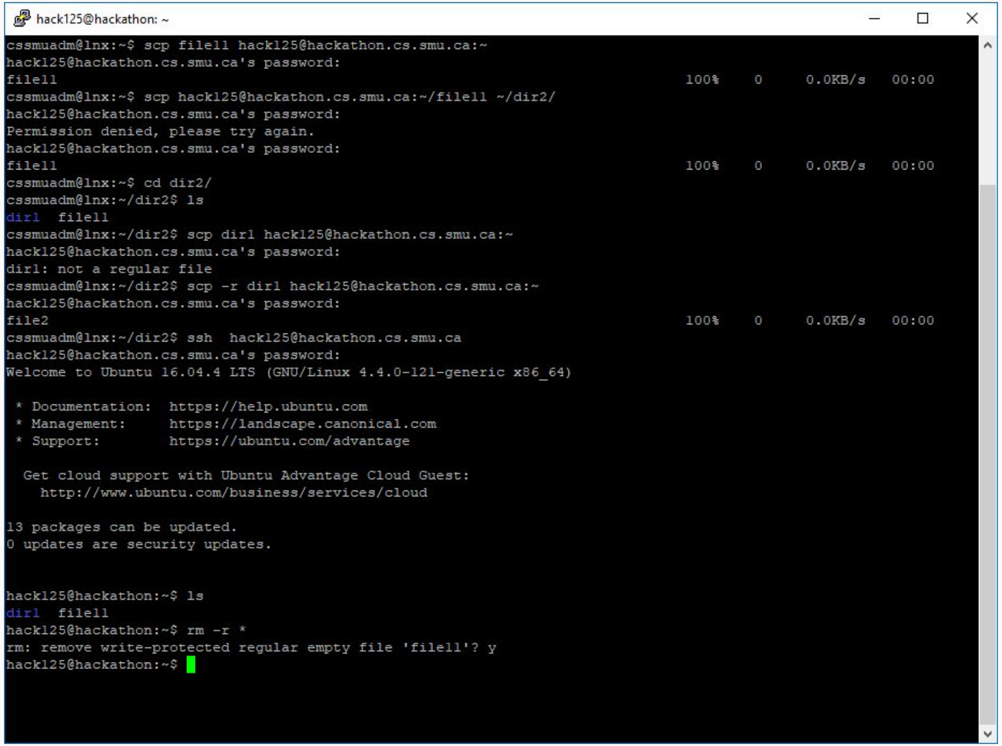
Use it like cp

scp   file   username@server:DIR

OR

scp   username@server:DIR   /TARGET_DIR

USE FLAG -r FOR COPYING FOLDERS



```
cssmuadm@lnx:~$ scp file11 hack125@hackathon.cs.smu.ca:~
hack125@hackathon.cs.smu.ca's password:
file11                                                                         100%    0     0.0KB/s   00:00
cssmuadm@lnx:~$ scp hack125@hackathon.cs.smu.ca:~/file11 ~/dir2/
hack125@hackathon.cs.smu.ca's password:
Permission denied, please try again.
hack125@hackathon.cs.smu.ca's password:
file11                                                                         100%    0     0.0KB/s   00:00
cssmuadm@lnx:~$ cd dir2/
cssmuadm@lnx:~/dir2$ ls
dir1  file11
cssmuadm@lnx:~/dir2$ scp dir1 hack125@hackathon.cs.smu.ca:~
hack125@hackathon.cs.smu.ca's password:
dir1: not a regular file
cssmuadm@lnx:~/dir2$ scp -r dir1 hack125@hackathon.cs.smu.ca:~
hack125@hackathon.cs.smu.ca's password:
file2                                                                          100%    0     0.0KB/s   00:00
cssmuadm@lnx:~/dir2$ ssh  hack125@hackathon.cs.smu.ca
hack125@hackathon.cs.smu.ca's password:
Welcome to Ubuntu 16.04.4 LTS (GNU/Linux 4.4.0-121-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

  Get cloud support with Ubuntu Advantage Cloud Guest:
    http://www.ubuntu.com/business/services/cloud

13 packages can be updated.
0 updates are security updates.


hack125@hackathon:~$ ls
dir1  file11
hack125@hackathon:~$ rm -r *
rm: remove write-protected regular empty file 'file11'? y
hack125@hackathon:~$
```

File Transfer

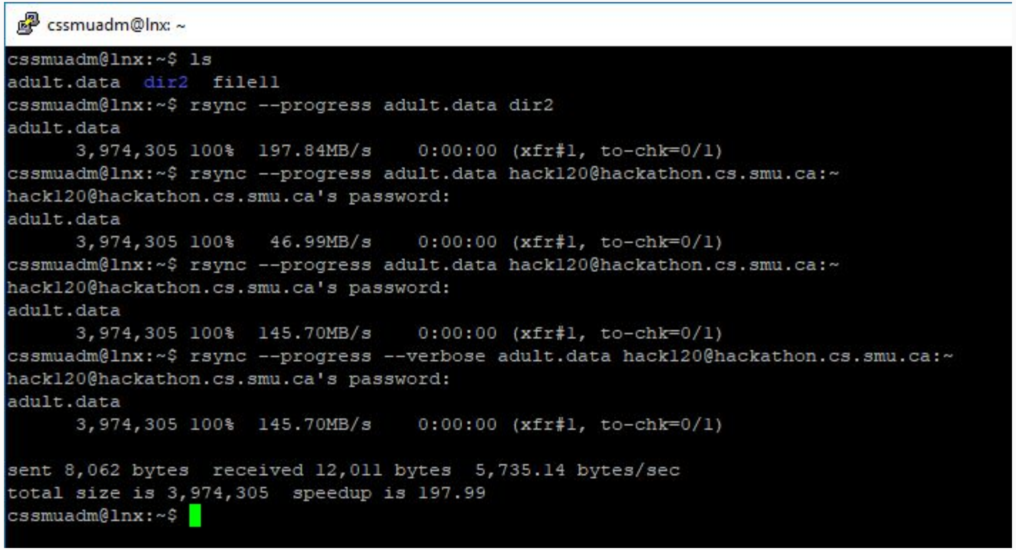
rsync

You can use it instead of both cp and scp. Also used for backups. Does not transmit unnecessary chunks as copies CHANGES (calculates checksums of blocks of file and compares them)

Useful flags:

- --progress
- --verbose

See man rsync for more details



```
cssmuadm@lnx:~$ ls
adult.data  dir2  file11
cssmuadm@lnx:~$ rsync --progress adult.data dir2
adult.data
      3,974,305 100%  197.84MB/s    0:00:00 (xfr#1, to-chk=0/1)
cssmuadm@lnx:~$ rsync --progress adult.data hack120@hackathon.cs.smu.ca:~
hack120@hackathon.cs.smu.ca's password:
adult.data
      3,974,305 100%   46.99MB/s    0:00:00 (xfr#1, to-chk=0/1)
cssmuadm@lnx:~$ rsync --progress adult.data hack120@hackathon.cs.smu.ca:~
hack120@hackathon.cs.smu.ca's password:
adult.data
      3,974,305 100%  145.70MB/s    0:00:00 (xfr#1, to-chk=0/1)
cssmuadm@lnx:~$ rsync --progress --verbose adult.data hack120@hackathon.cs.smu.ca:~
hack120@hackathon.cs.smu.ca's password:
adult.data
      3,974,305 100%  145.70MB/s    0:00:00 (xfr#1, to-chk=0/1)

sent 8,062 bytes  received 12,011 bytes  5,735.14 bytes/sec
total size is 3,974,305  speedup is 197.99
cssmuadm@lnx:~$
```

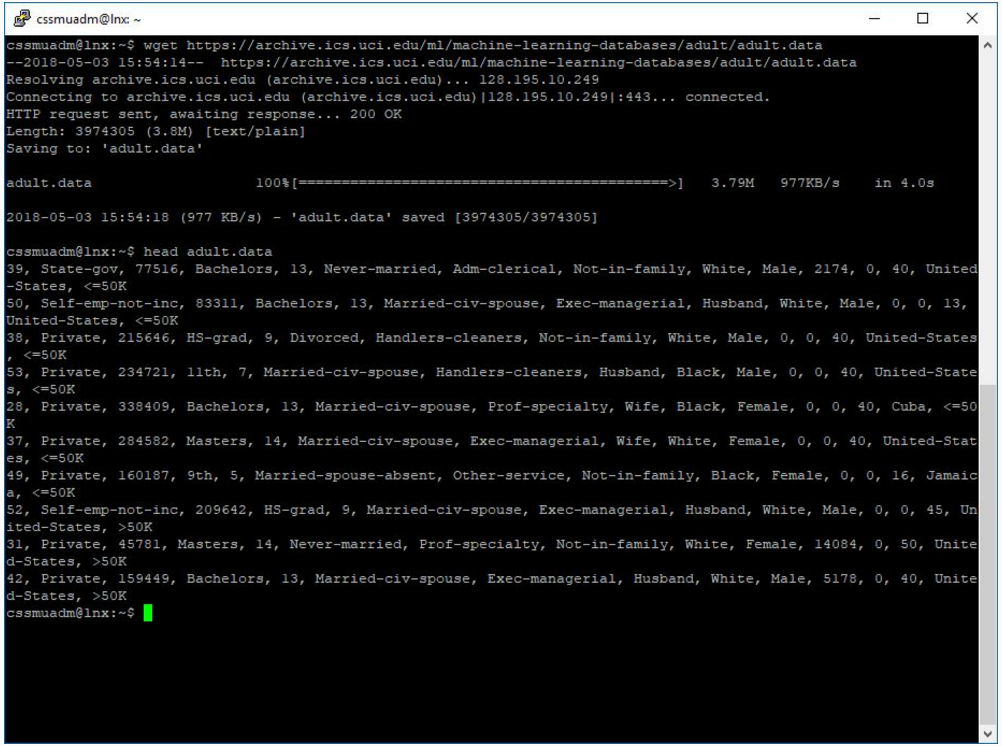
File Transfer

wget

wget link

Instead of downloading some files to your PC first then to server, just use wget where possible

See man wget for options/flags. But usually works fine without flags



```
cssmuadm@lnx:~$ wget https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data
--2018-05-03 15:54:14--  https://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.data
Resolving archive.ics.uci.edu (archive.ics.uci.edu)... 128.195.10.249
Connecting to archive.ics.uci.edu (archive.ics.uci.edu)|128.195.10.249|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 3974305 (3.8M) [text/plain]
Saving to: 'adult.data'

adult.data                 100%[===========================================>]   3.79M   977KB/s    in 4.0s

2018-05-03 15:54:18 (977 KB/s) - 'adult.data' saved [3974305/3974305]

cssmuadm@lnx:~$ head adult.data
39, State-gov, 77516, Bachelors, 13, Never-married, Adm-clerical, Not-in-family, White, Male, 2174, 0, 40, United
-States, <=50K
50, Self-emp-not-inc, 83311, Bachelors, 13, Married-civ-spouse, Exec-managerial, Husband, White, Male, 0, 0, 13,
United-States, <=50K
38, Private, 215646, HS-grad, 9, Divorced, Handlers-cleaners, Not-in-family, White, Male, 0, 0, 40, United-States
, <=50K
53, Private, 234721, 11th, 7, Married-civ-spouse, Handlers-cleaners, Husband, Black, Male, 0, 0, 40, United-State
s, <=50K
28, Private, 338409, Bachelors, 13, Married-civ-spouse, Prof-specialty, Wife, Black, Female, 0, 0, 40, Cuba, <=50
K
37, Private, 284582, Masters, 14, Married-civ-spouse, Exec-managerial, Wife, White, Female, 0, 0, 40, United-Stat
es, <=50K
49, Private, 160187, 9th, 5, Married-spouse-absent, Other-service, Not-in-family, Black, Female, 0, 0, 16, Jamaic
a, <=50K
52, Self-emp-not-inc, 209642, HS-grad, 9, Married-civ-spouse, Exec-managerial, Husband, White, Male, 0, 0, 45, Un
ited-States, >50K
31, Private, 45781, Masters, 14, Never-married, Prof-specialty, Not-in-family, White, Female, 14084, 0, 50, Unite
d-States, >50K
42, Private, 159449, Bachelors, 13, Married-civ-spouse, Exec-managerial, Husband, White, Male, 5178, 0, 40, Unite
d-States, >50K
cssmuadm@lnx:~$ 
```

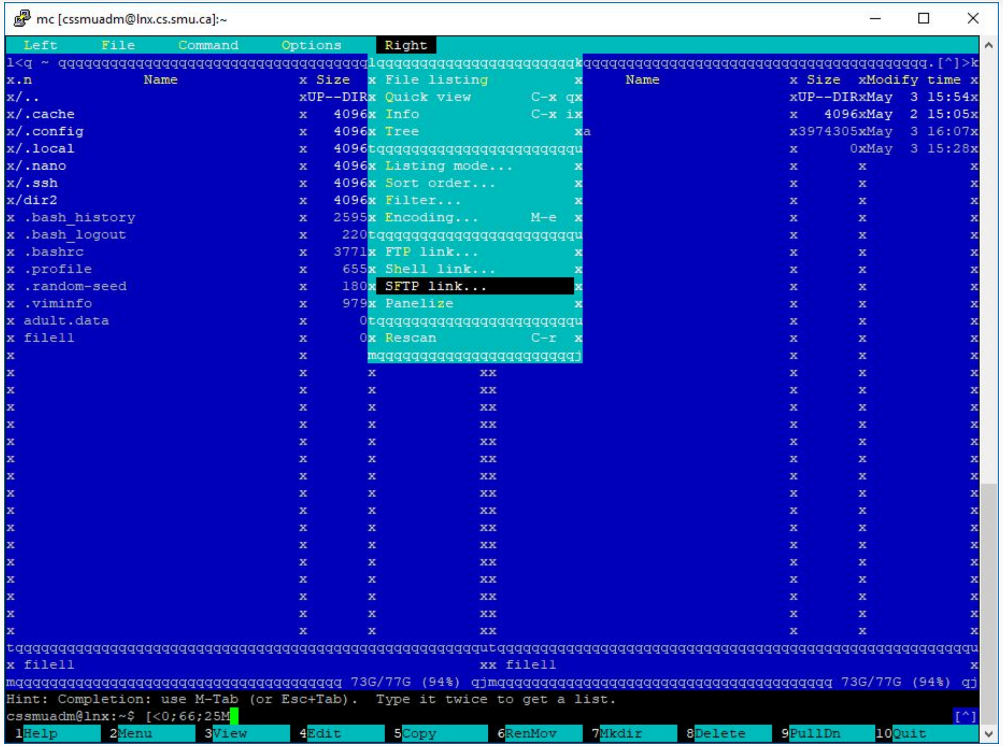
mc

mc - for those who do not feel comfortable with command line at all.

DOS-style file manager (like NC, or Total Commander), very simple text editor (mcedit), etc.

Supports SFTP as well

May not be installed on some systems



mc [cssmuadm@lnx.cs.smu.ca]:~

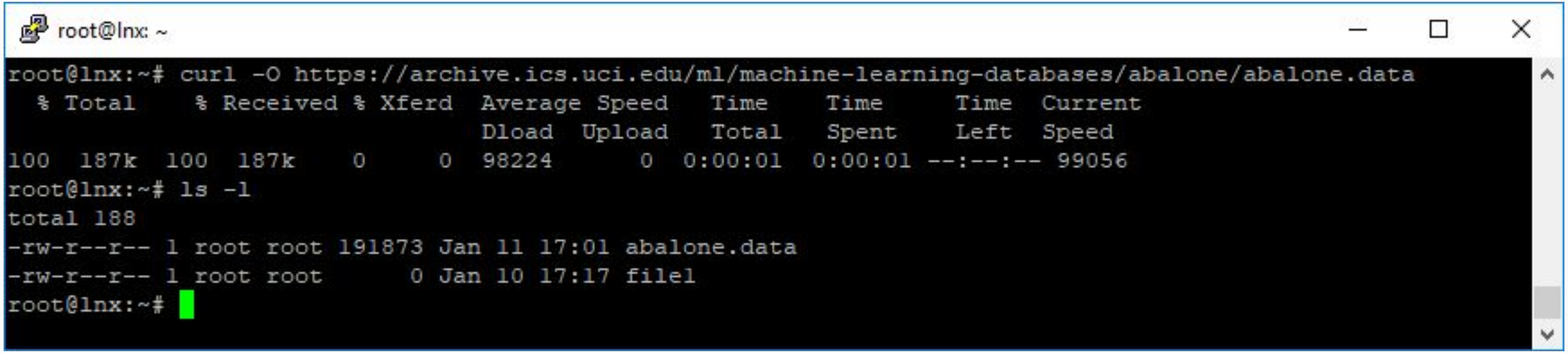
```
  Left      File      Command      Options     Right
l<q ~ qqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqlqqqqqqqqqqqqqqqqqqqqqqqqkqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq.[^]>k
x.n              Name               x Size x File listing          x       Name                x Size  xModify time x
x/..                                xUP--DIRx Quick view      C-x qx                           xUP--DIRxMay  3 15:54x
x/.cache                            x   4096x Info            C-x ix                           x   4096xMay  2 15:05x
x/.config                           x   4096x Tree                 xa                          x3974305xMay  3 16:07x
x/.local                            x   4096tqqqqqqqqqqqqqqqqqqqqqqqqu                         x       0xMay  3 15:28x
x/.nano                             x   4096x Listing mode...      x                           x       x            x
x/.ssh                              x   4096x Sort order...        x                           x       x            x
x/dir2                              x   4096x Filter...            x                           x       x            x
x .bash_history                     x   2595x Encoding...     M-e  x                           x       x            x
x .bash_logout                      x    220tqqqqqqqqqqqqqqqqqqqqqqqqu                         x       x            x
x .bashrc                           x   3771x FTP link...          x                           x       x            x
x .profile                          x    655x Shell link...        x                           x       x            x
x .random-seed                      x    180x SFTP link...         x                           x       x            x
x .viminfo                          x    979x Panelize             x                           x       x            x
x adult.data                        x      0tqqqqqqqqqqqqqqqqqqqqqqqqu                         x       x            x
x filell                            x      0x Rescan          C-r  x                           x       x            x
x                                   x       mqqqqqqqqqqqqqqqqqqqqqqqqj                         x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
x                                   x       x             xx                                   x       x            x
tqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqutqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqu
x filell                                                  xx filell                                                x
mqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq 73G/77G (94%) qjmqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq 73G/77G (94%) qj
Hint: Completion: use M-Tab (or Esc+Tab).  Type it twice to get a list.
cssmuadm@lnx:~$ [<0;66;25M                                                                                   [^]
 1Help     2Menu     3View     4Edit     5Copy      6RenMov    7Mkdir    8Delete   9PullDn   10Quit
```

Exercise

1. Download any dataset from https://archive.ics.uci.edu/ml/datasets.html using wget to your home folder at lnx.cs.smu.ca (e.g., https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.data)
2. Copy it from lnx.cs.smu.ca to your desktop using WinSCP (or other software)
3. Copy it from lnx.cs.smu.ca to dev.cs.smu.ca using either scp or rsync command
4. ssh to dev.cs.smu.ca from lnx.cs.smu.ca and check if the file is there

curl

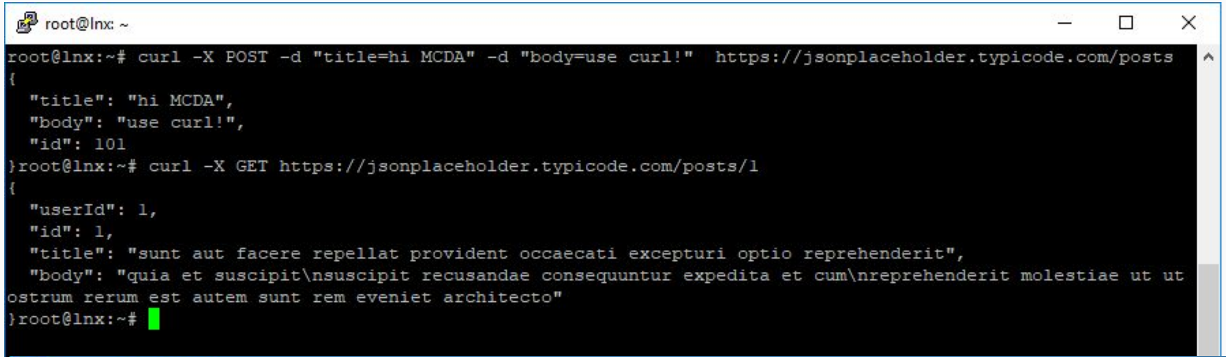
curl - can be used for downloading files as well with -O option



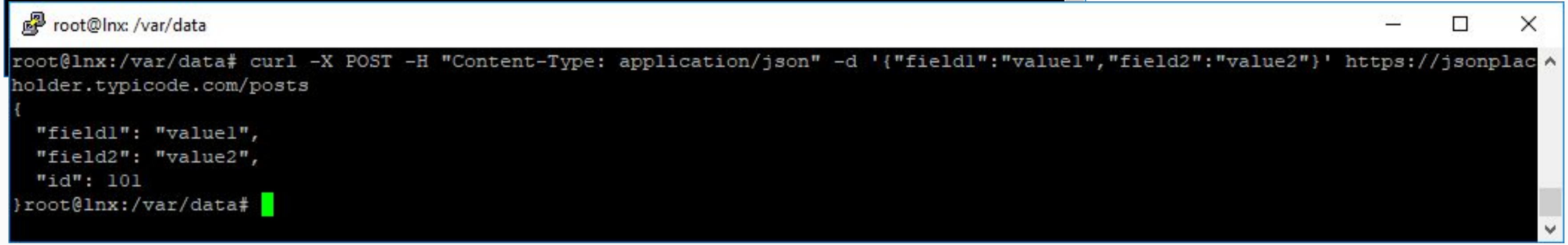
```
root@lnx:~# curl -O https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.data
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100  187k  100  187k    0     0  98224      0  0:00:01  0:00:01 --:--:-- 99056
root@lnx:~# ls -l
total 188
-rw-r--r-- 1 root root 191873 Jan 11 17:01 abalone.data
-rw-r--r-- 1 root root      0 Jan 10 17:17 file1
root@lnx:~# 
```

curl again

curl - easy way to test RESTful web services (supports various HTTP methods)



```
root@lnx:~# curl -X POST -d "title=hi MCDA" -d "body=use curl!"  https://jsonplaceholder.typicode.com/posts
{
  "title": "hi MCDA",
  "body": "use curl!",
  "id": 101
}root@lnx:~# curl -X GET https://jsonplaceholder.typicode.com/posts/1
{
  "userId": 1,
  "id": 1,
  "title": "sunt aut facere repellat provident occaecati excepturi optio reprehenderit",
  "body": "quia et suscipit\nsuscipit recusandae consequuntur expedita et cum\nreprehenderit molestiae ut ut
ostrum rerum est autem sunt rem eveniet architecto"
}root@lnx:~# 
```



```
root@lnx:/var/data# curl -X POST -H "Content-Type: application/json" -d '{"field1":"value1","field2":"value2"}' https://jsonplac
holder.typicode.com/posts
{
  "field1": "value1",
  "field2": "value2",
  "id": 101
}root@lnx:/var/data# 
```

Exercise

Send POST request to https://jsonplaceholder.typicode.com/posts using curl with the following fields both in Json format and as simple POST parameters:

- fname: your first name
- lname: your last name
- course: MCDA5570

Archivers

3 types of archives:

- Archive only with no compression (e.g. tar, cpio, ar)
- Compress only (e.g., bzip2, gzip, lzip)
- Compress + archive (7z, zip, rar, jar, war, etc.)

Archivers

tar (tarballs)

- DOES NOT COMPRESS FILES, just “collects” them to 1 file - tarball

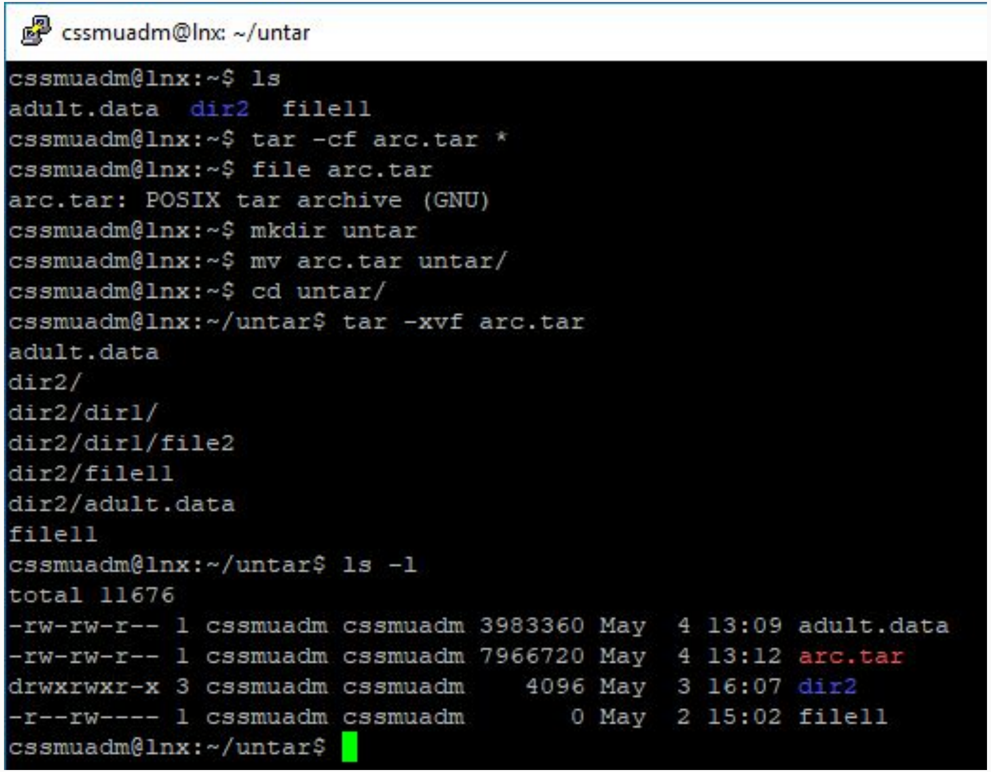
[link]

To archive files:

- tar -cf arc.tar file1 file2 file3
  - Where -c means COMPRESS
  - f means filename of archive

To extract files:

- tar -xvf arc.tar
  - Where -x means eXtract
  - f means tarball Filename
  - v means Verbose (optional)



```
cssmuadm@lnx:~$ ls
adult.data  dir2  file11
cssmuadm@lnx:~$ tar -cf arc.tar *
cssmuadm@lnx:~$ file arc.tar
arc.tar: POSIX tar archive (GNU)
cssmuadm@lnx:~$ mkdir untar
cssmuadm@lnx:~$ mv arc.tar untar/
cssmuadm@lnx:~$ cd untar/
cssmuadm@lnx:~/untar$ tar -xvf arc.tar
adult.data
dir2/
dir2/dir1/
dir2/dir1/file2
dir2/file11
dir2/adult.data
file11
cssmuadm@lnx:~/untar$ ls -l
total 11676
-rw-rw-r-- 1 cssmuadm cssmuadm 3983360 May  4 13:09 adult.data
-rw-rw-r-- 1 cssmuadm cssmuadm 7966720 May  4 13:12 arc.tar
drwxrwxr-x 3 cssmuadm cssmuadm    4096 May  3 16:07 dir2
-r--rw---- 1 cssmuadm cssmuadm       0 May  2 15:02 file11
cssmuadm@lnx:~/untar$ 
```

Archivers

gzip

- Compresses single files using “deflate” algorithm

To compress files:

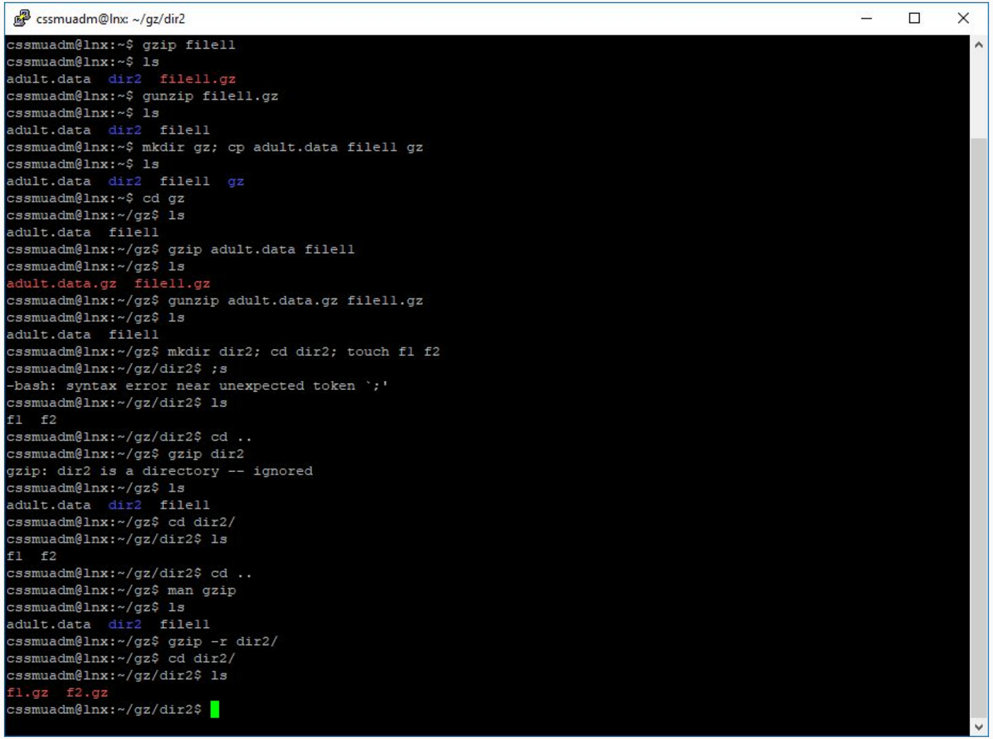
- gzip file1

To extract files:

- gunzip file1.gz

Useful flags: -r : recursively

!Not convenient to compress/extract separate files!



```
cssmuadm@lnx:~$ gzip file11
cssmuadm@lnx:~$ ls
adult.data  dir2  file11.gz
cssmuadm@lnx:~$ gunzip file11.gz
cssmuadm@lnx:~$ ls
adult.data  dir2  file11
cssmuadm@lnx:~$ mkdir gz; cp adult.data file11 gz
cssmuadm@lnx:~$ ls
adult.data  dir2  file11  gz
cssmuadm@lnx:~$ cd gz
cssmuadm@lnx:~/gz$ ls
adult.data  file11
cssmuadm@lnx:~/gz$ gzip adult.data file11
cssmuadm@lnx:~/gz$ ls
adult.data.gz  file11.gz
cssmuadm@lnx:~/gz$ gunzip adult.data.gz file11.gz
cssmuadm@lnx:~/gz$ ls
adult.data  file11
cssmuadm@lnx:~/gz$ mkdir dir2; cd dir2; touch f1 f2
cssmuadm@lnx:~/gz/dir2$ ;s
-bash: syntax error near unexpected token `;'
cssmuadm@lnx:~/gz/dir2$ ls
f1  f2
cssmuadm@lnx:~/gz/dir2$ cd ..
cssmuadm@lnx:~/gz$ gzip dir2
gzip: dir2 is a directory -- ignored
cssmuadm@lnx:~/gz$ ls
adult.data  dir2  file11
cssmuadm@lnx:~/gz$ cd dir2/
cssmuadm@lnx:~/gz/dir2$ ls
f1  f2
cssmuadm@lnx:~/gz/dir2$ cd ..
cssmuadm@lnx:~/gz$ man gzip
cssmuadm@lnx:~/gz$ ls
adult.data  dir2  file11
cssmuadm@lnx:~/gz$ gzip -r dir2/
cssmuadm@lnx:~/gz$ cd dir2/
cssmuadm@lnx:~/gz/dir2$ ls
f1.gz  f2.gz
cssmuadm@lnx:~/gz/dir2$ 
```

Archivers

tar + gzip

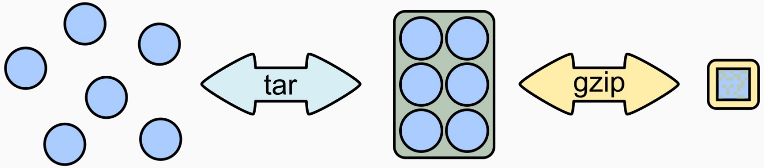
To archive files with gzip compression:

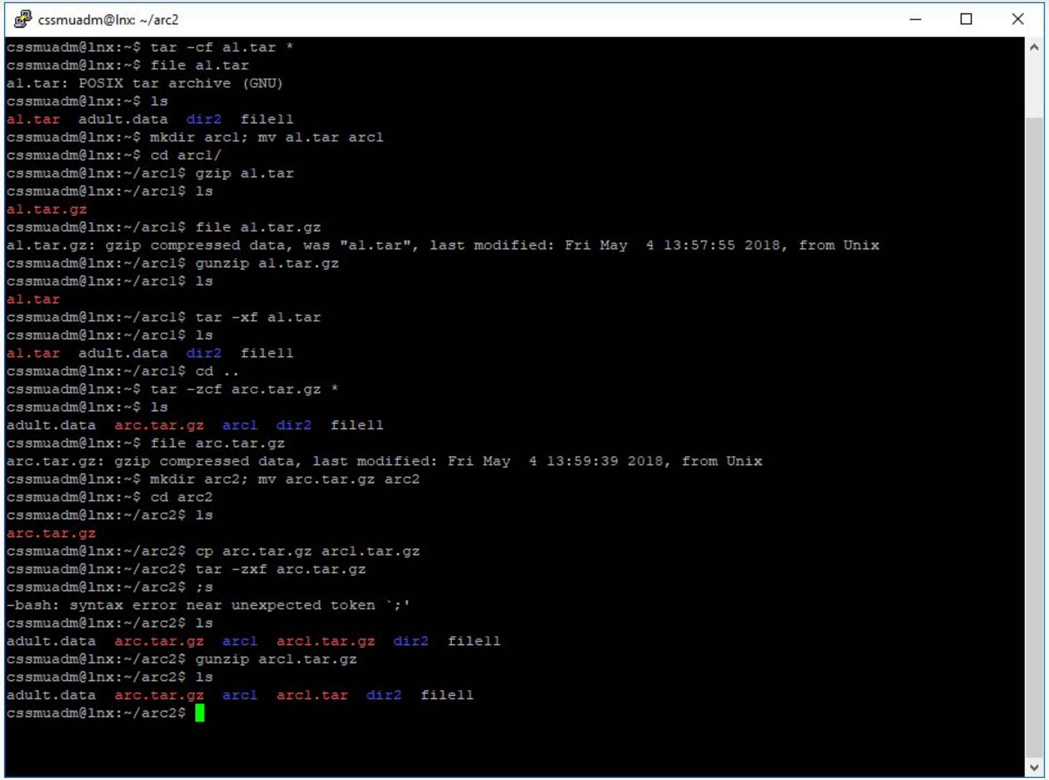
- tar -zcf file1 file2 file3
    - Where -z means compress using gzip

To extract files:

- tar -zxf arc.tgz

.tgz = .tar.gz





```
cssmuadm@lnx:~$ tar -cf al.tar *
cssmuadm@lnx:~$ file al.tar
al.tar: POSIX tar archive (GNU)
cssmuadm@lnx:~$ ls
al.tar  adult.data  dir2  filell
cssmuadm@lnx:~$ mkdir arcl; mv al.tar arcl
cssmuadm@lnx:~$ cd arcl/
cssmuadm@lnx:~/arcl$ gzip al.tar
cssmuadm@lnx:~/arcl$ ls
al.tar.gz
cssmuadm@lnx:~/arcl$ file al.tar.gz
al.tar.gz: gzip compressed data, was "al.tar", last modified: Fri May  4 13:57:55 2018, from Unix
cssmuadm@lnx:~/arcl$ gunzip al.tar.gz
cssmuadm@lnx:~/arcl$ ls
al.tar
cssmuadm@lnx:~/arcl$ tar -xf al.tar
cssmuadm@lnx:~/arcl$ ls
al.tar  adult.data  dir2  filell
cssmuadm@lnx:~/arcl$ cd ..
cssmuadm@lnx:~$ tar -zcf arc.tar.gz *
cssmuadm@lnx:~$ ls
adult.data  arc.tar.gz  arcl  dir2  filell
cssmuadm@lnx:~$ file arc.tar.gz
arc.tar.gz: gzip compressed data, last modified: Fri May  4 13:59:39 2018, from Unix
cssmuadm@lnx:~$ mkdir arc2; mv arc.tar.gz arc2
cssmuadm@lnx:~$ cd arc2
cssmuadm@lnx:~/arc2$ ls
arc.tar.gz
cssmuadm@lnx:~/arc2$ cp arc.tar.gz arcl.tar.gz
cssmuadm@lnx:~/arc2$ tar -zxf arc.tar.gz
cssmuadm@lnx:~/arc2$ ;s
-bash: syntax error near unexpected token `;'
cssmuadm@lnx:~/arc2$ ls
adult.data  arc.tar.gz  arcl  arcl.tar.gz  dir2  filell
cssmuadm@lnx:~/arc2$ gunzip arcl.tar.gz
cssmuadm@lnx:~/arc2$ ls
adult.data  arc.tar.gz  arcl  arcl.tar  dir2  filell
cssmuadm@lnx:~/arc2$ 
```

Archivers

bzip2

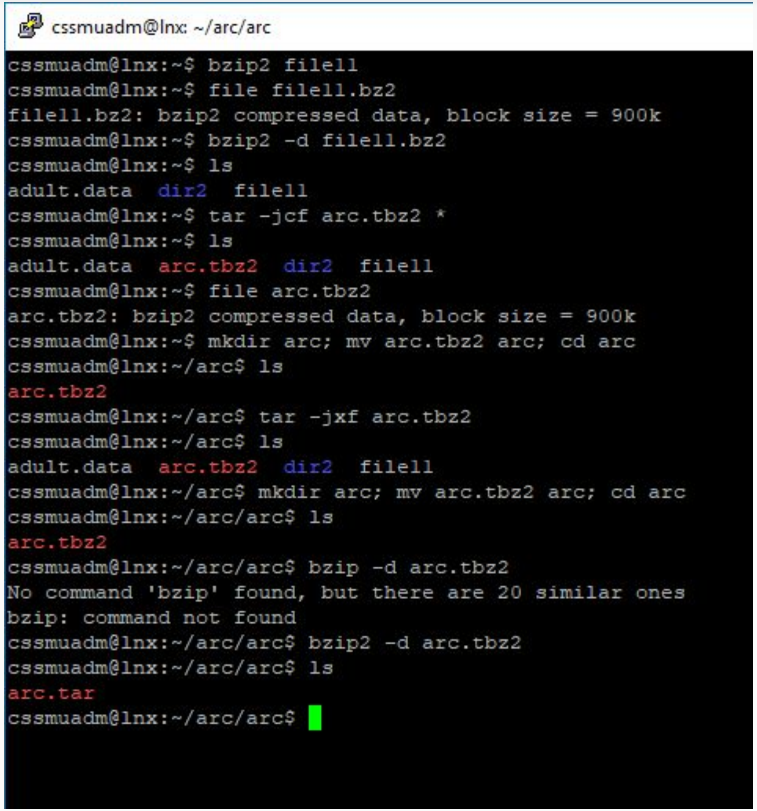
Uses Burrows–Wheeler transform for compression

- bzip2 file1 (compress file)
- bzip2 -d file1.xz (extract file with -d flag)

With tar:

- tar -jcf arc.tbz2 file1 file2 file2
  - Where -j means compress files with bzip2
- tar -jxf arc.tbz2

.tbz2 = .tar.bzip2 = .tbz = .tb2



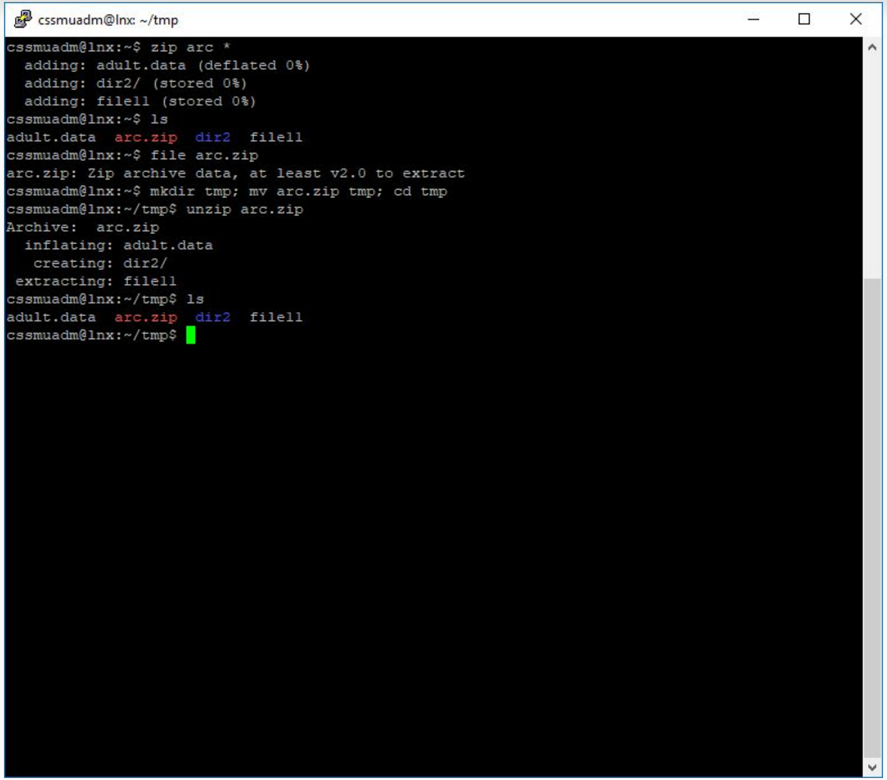
```
cssmuadm@lnx:~$ bzip2 file11
cssmuadm@lnx:~$ file file11.bz2
file11.bz2: bzip2 compressed data, block size = 900k
cssmuadm@lnx:~$ bzip2 -d file11.bz2
cssmuadm@lnx:~$ ls
adult.data  dir2  file11
cssmuadm@lnx:~$ tar -jcf arc.tbz2 *
cssmuadm@lnx:~$ ls
adult.data  arc.tbz2  dir2  file11
cssmuadm@lnx:~$ file arc.tbz2
arc.tbz2: bzip2 compressed data, block size = 900k
cssmuadm@lnx:~$ mkdir arc; mv arc.tbz2 arc; cd arc
cssmuadm@lnx:~/arc$ ls
arc.tbz2
cssmuadm@lnx:~/arc$ tar -jxf arc.tbz2
cssmuadm@lnx:~/arc$ ls
adult.data  arc.tbz2  dir2  file11
cssmuadm@lnx:~/arc$ mkdir arc; mv arc.tbz2 arc; cd arc
cssmuadm@lnx:~/arc/arc$ ls
arc.tbz2
cssmuadm@lnx:~/arc/arc$ bzip -d arc.tbz2
No command 'bzip' found, but there are 20 similar ones
bzip: command not found
cssmuadm@lnx:~/arc/arc$ bzip2 -d arc.tbz2
cssmuadm@lnx:~/arc/arc$ ls
arc.tar
cssmuadm@lnx:~/arc/arc$
```

Archivers

zip

Uses different algorithms to compress files (usually “deflate” as gzip).

- zip arc file1 file2 file3 (compress and archive files, where arc is a name of archive)
- unzip arc .zip (extract files)



```
cssmuadm@lnx:~$ zip arc *
  adding: adult.data (deflated 0%)
  adding: dir2/ (stored 0%)
  adding: filell (stored 0%)
cssmuadm@lnx:~$ ls
adult.data  arc.zip  dir2  filell
cssmuadm@lnx:~$ file arc.zip
arc.zip: Zip archive data, at least v2.0 to extract
cssmuadm@lnx:~$ mkdir tmp; mv arc.zip tmp; cd tmp
cssmuadm@lnx:~/tmp$ unzip arc.zip
Archive:  arc.zip
  inflating: adult.data
   creating: dir2/
 extracting: filell
cssmuadm@lnx:~/tmp$ ls
adult.data  arc.zip  dir2  filell
cssmuadm@lnx:~/tmp$
```

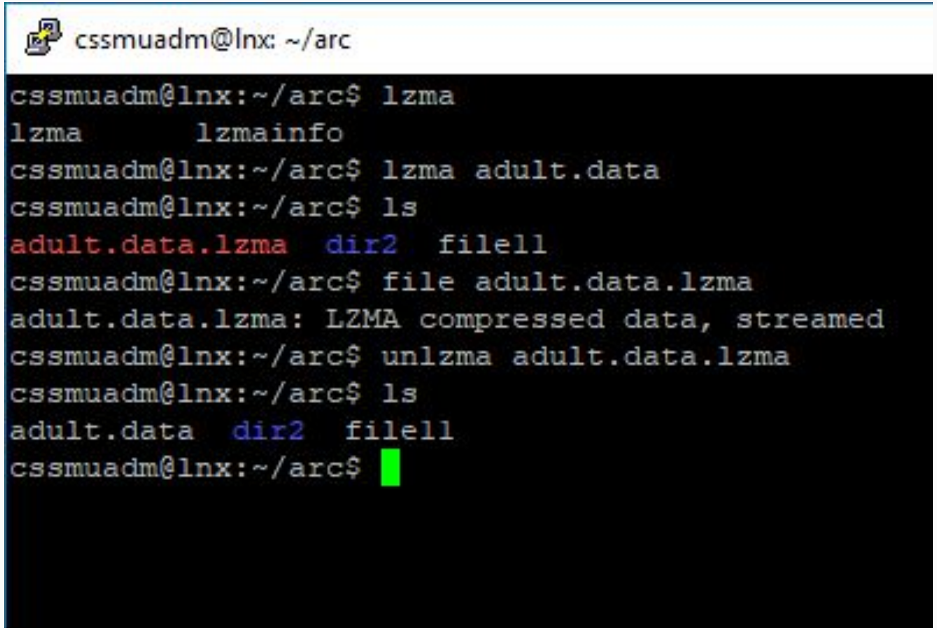
Archivers

lzma

Uses Lempel–Ziv–Markov chain algorithm to compress files (as 7zip).

- lzma file1 (compress file)
- unlzma file1.lzma (extract file)

.tbz2 = .tbz = .tar.lzma = .tlz



```
cssmuadm@lnx:~/arc$ lzma
lzma       lzmainfo
cssmuadm@lnx:~/arc$ lzma adult.data
cssmuadm@lnx:~/arc$ ls
adult.data.lzma  dir2  filell
cssmuadm@lnx:~/arc$ file adult.data.lzma
adult.data.lzma: LZMA compressed data, streamed
cssmuadm@lnx:~/arc$ unlzma adult.data.lzma
cssmuadm@lnx:~/arc$ ls
adult.data  dir2  filell
cssmuadm@lnx:~/arc$
```

Archivers

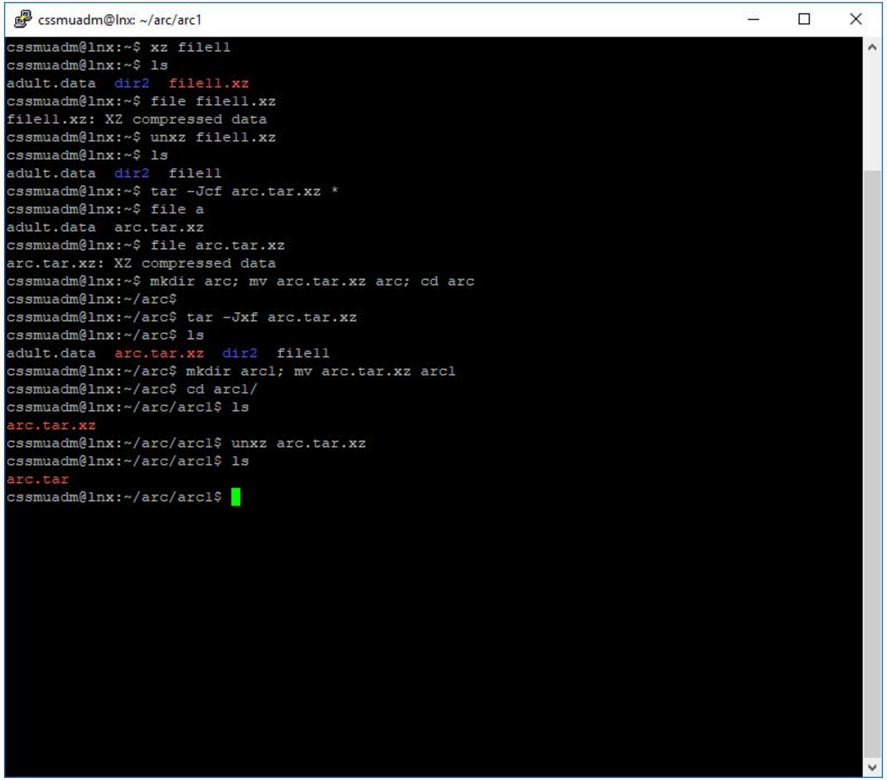
xz

xz is a compression program and file format. Uses LZMA/LZMA2 compression algorithms.

- xz file1 (compress file)
- unxz file1.xz (extract file)

With tar:

- tar -Jcf arc.tar.xz file1 file2 file2
  - Where -J means compress files with xz
- tar -Jxf arc.tar.xz



```
cssmuadm@lnx:~$ xz filel1
cssmuadm@lnx:~$ ls
adult.data  dir2  filel1.xz
cssmuadm@lnx:~$ file filel1.xz
filel1.xz: XZ compressed data
cssmuadm@lnx:~$ unxz filel1.xz
cssmuadm@lnx:~$ ls
adult.data  dir2  filel1
cssmuadm@lnx:~$ tar -Jcf arc.tar.xz *
cssmuadm@lnx:~$ file a
adult.data  arc.tar.xz
cssmuadm@lnx:~$ file arc.tar.xz
arc.tar.xz: XZ compressed data
cssmuadm@lnx:~$ mkdir arc; mv arc.tar.xz arc; cd arc
cssmuadm@lnx:~/arc$
cssmuadm@lnx:~/arc$ tar -Jxf arc.tar.xz
cssmuadm@lnx:~/arc$ ls
adult.data  arc.tar.xz  dir2  filel1
cssmuadm@lnx:~/arc$ mkdir arc1; mv arc.tar.xz arc1
cssmuadm@lnx:~/arc$ cd arc1/
cssmuadm@lnx:~/arc/arc1$ ls
arc.tar.xz
cssmuadm@lnx:~/arc/arc1$ unxz arc.tar.xz
cssmuadm@lnx:~/arc/arc1$ ls
arc.tar
cssmuadm@lnx:~/arc/arc1$
```

Archivers

Miscellaneous:

- For .7zip use p7zip (usually not a part of default set of software)
- For .jar and .war use jar
- For .rar use unrar (usually not a part of default set of software)

Exercise

Two groups: one group compresses binary files, another group - text files

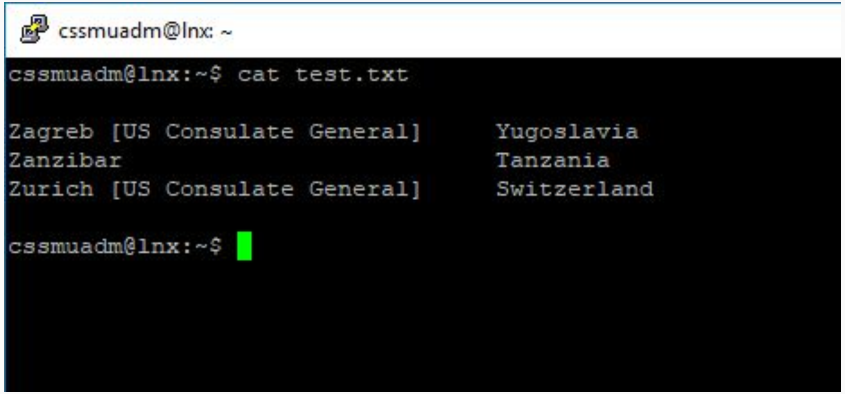
1. Compress /bin (GROUP 1) and /var/data (GROUP 2) folder with tar using different flags for archiving (j/J/z)
2. Compare sizes of resulting archives. Which option gives better compression?

Working with text data

cat

Outputs contents of text file(s) to terminal.

Good for small files. Do not use with large files or binary files



```
cssmuadm@lnx:~$ cat test.txt

Zagreb [US Consulate General]     Yugoslavia
Zanzibar                          Tanzania
Zurich [US Consulate General]     Switzerland

cssmuadm@lnx:~$ 
```

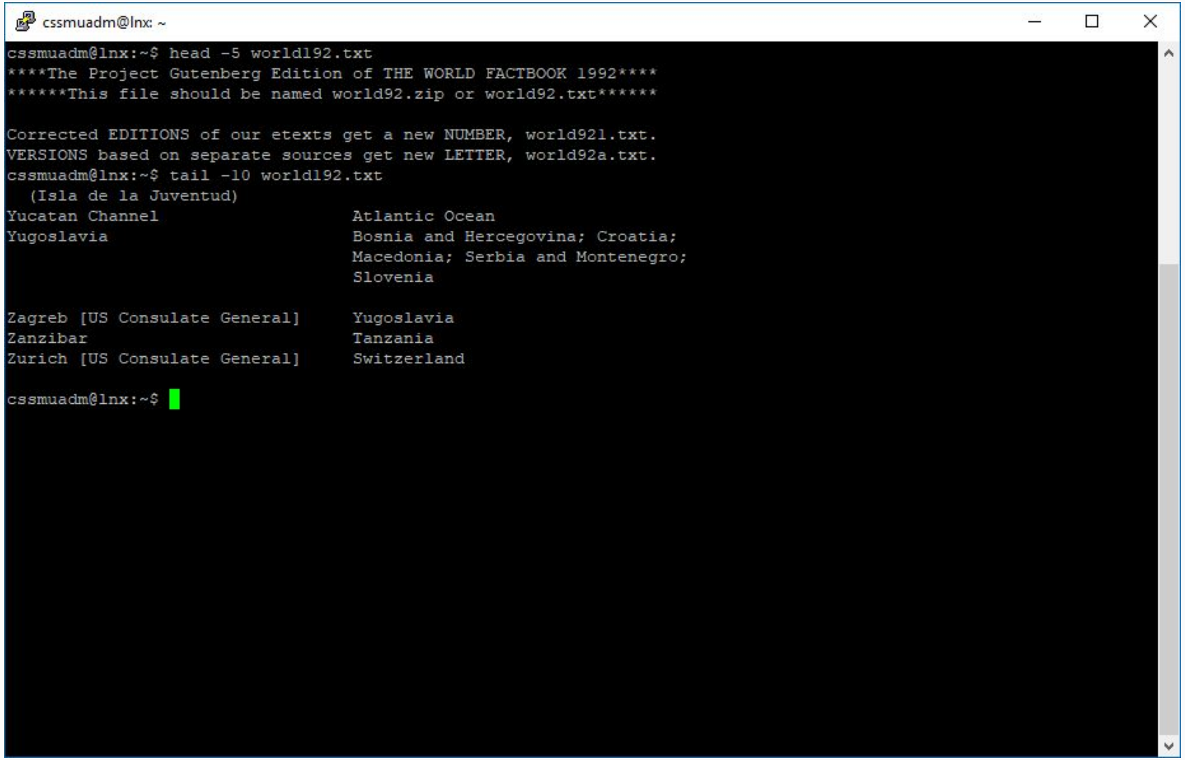
[link]

Working with text data

head/tail

output the first/last part of text files

Number of lines to output may be specified after hyphen (e.g., - 10)



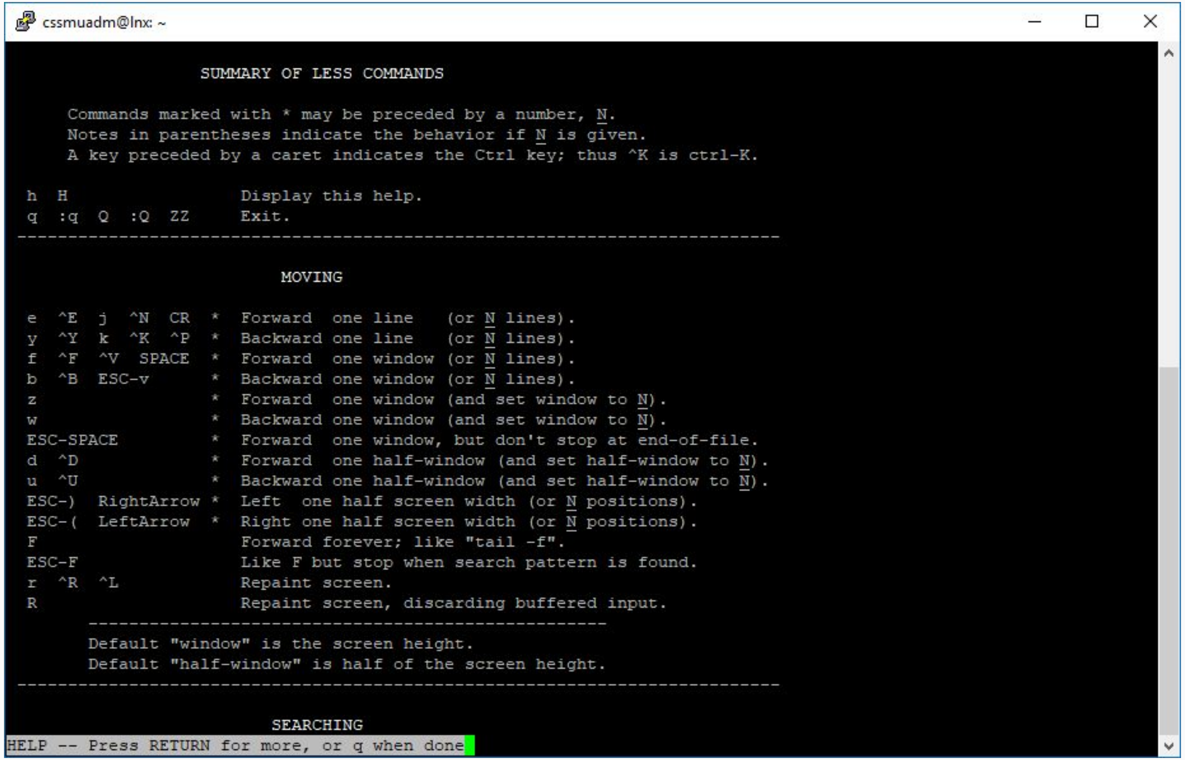
```
cssmuadm@lnx:~$ head -5 world192.txt
****The Project Gutenberg Edition of THE WORLD FACTBOOK 1992****
******This file should be named world92.zip or world92.txt******

Corrected EDITIONS of our etexts get a new NUMBER, world921.txt.
VERSIONS based on separate sources get new LETTER, world92a.txt.
cssmuadm@lnx:~$ tail -10 world192.txt
  (Isla de la Juventud)
Yucatan Channel                   Atlantic Ocean
Yugoslavia                        Bosnia and Hercegovina; Croatia;
                                  Macedonia; Serbia and Montenegro;
                                  Slovenia

Zagreb [US Consulate General]     Yugoslavia
Zanzibar                          Tanzania
Zurich [US Consulate General]     Switzerland

cssmuadm@lnx:~$
```

Working with text data

less

less textfile

View large files. PgUp/PgDown for scrolling + additional commands (see help)

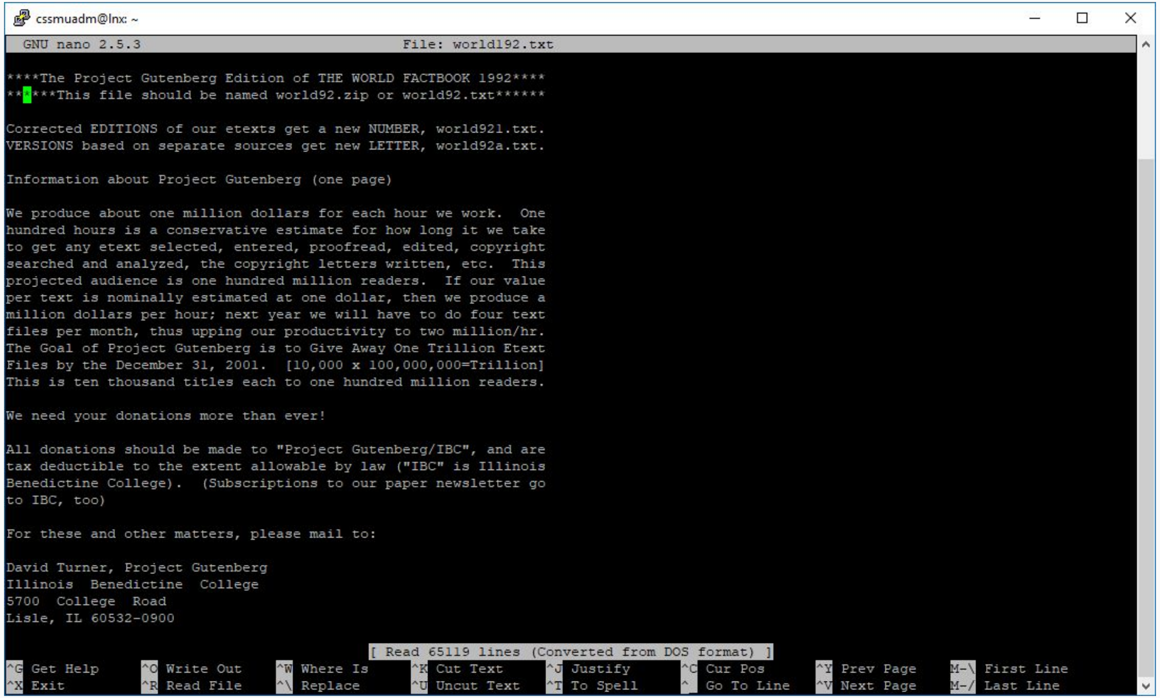


```
                   SUMMARY OF LESS COMMANDS

      Commands marked with * may be preceded by a number, N.
      Notes in parentheses indicate the behavior if N is given.
      A key preceded by a caret indicates the Ctrl key; thus ^K is ctrl-K.

  h  H                 Display this help.
  q  :q  Q  :Q  ZZ     Exit.
 ---------------------------------------------------------------------------

                           MOVING

  e  ^E  j  ^N  CR  *  Forward  one line   (or N lines).
  y  ^Y  k  ^K  ^P  *  Backward one line   (or N lines).
  f  ^F  ^V  SPACE  *  Forward  one window (or N lines).
  b  ^B  ESC-v      *  Backward one window (or N lines).
  z                 *  Forward  one window (and set window to N).
  w                 *  Backward one window (and set window to N).
  ESC-SPACE         *  Forward  one window, but don't stop at end-of-file.
  d  ^D             *  Forward  one half-window (and set half-window to N).
  u  ^U             *  Backward one half-window (and set half-window to N).
  ESC-)  RightArrow *  Left  one half screen width (or N positions).
  ESC-(  LeftArrow  *  Right one half screen width (or N positions).
  F                    Forward forever; like "tail -f".
  ESC-F                Like F but stop when search pattern is found.
  r  ^R  ^L            Repaint screen.
  R                    Repaint screen, discarding buffered input.
        ---------------------------------------------------
        Default "window" is the screen height.
        Default "half-window" is half of the screen height.
 ---------------------------------------------------------------------------

                          SEARCHING
HELP -- Press RETURN for more, or q when done
```

Working with text data

nano (cmd: nano *text.txt*) is one of the simplest text editors.
Can copy and paste with mouse.

- Ctrl+X to exit
- Ctrl+S to save
- Ctrl+W to find
- Ctrl+K to cut
- Ctrl+U to paste



```
  GNU nano 2.5.3                              File: world192.txt

****The Project Gutenberg Edition of THE WORLD FACTBOOK 1992****
*****This file should be named world92.zip or world92.txt******

Corrected EDITIONS of our etexts get a new NUMBER, world921.txt.
VERSIONS based on separate sources get new LETTER, world92a.txt.

Information about Project Gutenberg (one page)

We produce about one million dollars for each hour we work.  One
hundred hours is a conservative estimate for how long it we take
to get any etext selected, entered, proofread, edited, copyright
searched and analyzed, the copyright letters written, etc.  This
projected audience is one hundred million readers.  If our value
per text is nominally estimated at one dollar, then we produce a
million dollars per hour; next year we will have to do four text
files per month, thus upping our productivity to two million/hr.
The Goal of Project Gutenberg is to Give Away One Trillion Etext
Files by the December 31, 2001.  [10,000 x 100,000,000=Trillion]
This is ten thousand titles each to one hundred million readers.

We need your donations more than ever!

All donations should be made to "Project Gutenberg/IBC", and are
tax deductible to the extent allowable by law ("IBC" is Illinois
Benedictine College).  (Subscriptions to our paper newsletter go
to IBC, too)

For these and other matters, please mail to:

David Turner, Project Gutenberg
Illinois  Benedictine  College
5700  College  Road
Lisle, IL 60532-0900

                                            [ Read 65119 lines (Converted from DOS format) ]
^G Get Help     ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify      ^C Cur Pos      ^Y Prev Page    M-\ First Line
^X Exit         ^R Read File    ^\ Replace      ^U Uncut Text   ^T To Spell     ^_ Go To Line   ^V Next Page    M-/ Last Line
```

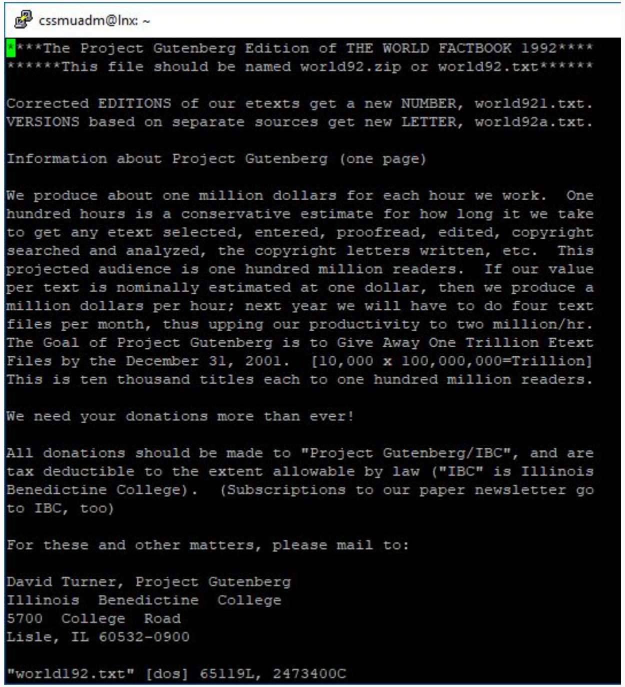
Working with text data

vim (cmd: vi *text.txt*) - powerful text editor. Great for source code.

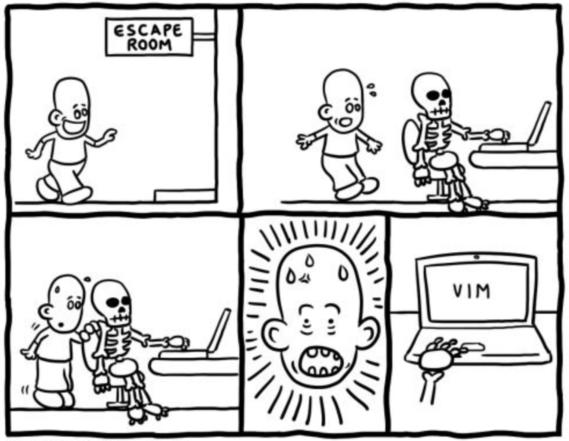
Has multiple MODES:

- i - insert mode. Can edit text almost like in nano
- Normal mode (press ESC from insert mode) - has navigation commands, save/exit, etc.

:q - exit, :w write, etc..



```
****The Project Gutenberg Edition of THE WORLD FACTBOOK 1992****
******This file should be named world92.zip or world92.txt******

Corrected EDITIONS of our etexts get a new NUMBER, world921.txt.
VERSIONS based on separate sources get new LETTER, world92a.txt.

Information about Project Gutenberg (one page)

We produce about one million dollars for each hour we work.  One
hundred hours is a conservative estimate for how long it we take
to get any etext selected, entered, proofread, edited, copyright
searched and analyzed, the copyright letters written, etc.  This
projected audience is one hundred million readers.  If our value
per text is nominally estimated at one dollar, then we produce a
million dollars per hour; next year we will have to do four text
files per month, thus upping our productivity to two million/hr.
The Goal of Project Gutenberg is to Give Away One Trillion Etext
Files by the December 31, 2001.  [10,000 x 100,000,000=Trillion]
This is ten thousand titles each to one hundred million readers.

We need your donations more than ever!

All donations should be made to "Project Gutenberg/IBC", and are
tax deductible to the extent allowable by law ("IBC" is Illinois
Benedictine College).  (Subscriptions to our paper newsletter go
to IBC, too)

For these and other matters, please mail to:

David Turner, Project Gutenberg
Illinois  Benedictine  College
5700  College  Road
Lisle, IL 60532-0900

"world192.txt" [dos] 65119L, 2473400C
```



Daniel Stori {turnoff.us}

He never knew how to quit...

[link]

Exercise

1. Copy any file from /var/data
2. Print first 5 and last 10 lines of the file with to console
3. Try to edit the file with either vim or nano