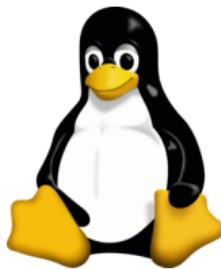


# Introduction to Linux

Maksim Kholmatov

EMBL

10 April 2024



# What is Linux?

## What does Wikipedia say?

Linux is a family of open-source **Unix-like** operating systems based on the Linux kernel, an operating system kernel first released on September 17, 1991, by Linus Torvalds.

## Unix-like?

A **Unix-like** (sometimes referred to as UN\*X or \*nix) operating system is one that behaves in a manner similar to a Unix system, while not necessarily conforming to or being certified to any version of the Single UNIX Specification. A Unix-like application is one that behaves like the corresponding Unix command or shell.

**NB!** - most of today's concepts will apply to most *Unix-like* operating systems.



# Time sharing systems



Figure 1: Ken Thompson and Dennis Ritchie working on a PDP-11



# Unix shell

A Unix shell is a **command-line interpreter** or shell that provides a command line user interface for Unix-like operating systems. The shell is both an interactive **command language** and a **scripting language**, and is used by the operating system to control the execution of the system using shell scripts.

## Commonly used shells

- Bourne Shell sh
- Bourne Again Shell bash
- Korn Shell ksh
- Z Shell zsh

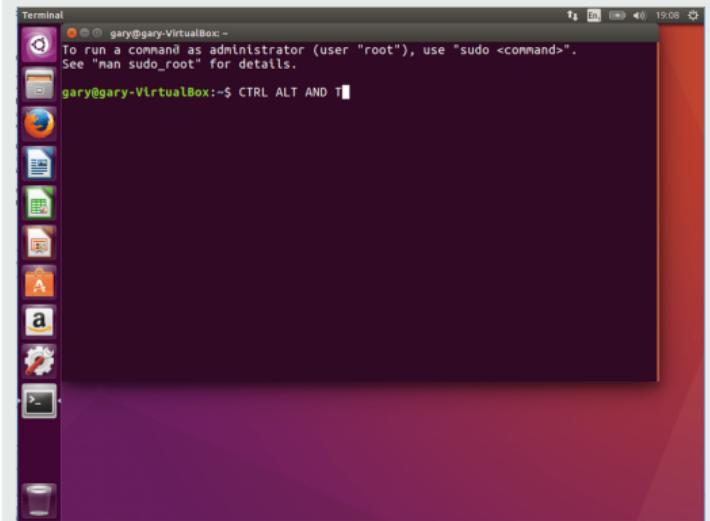
Users typically interact with a Unix shell using a **terminal emulator**; however, direct operation via serial hardware connections or Secure Shell ssh are common for server systems.

# Terminal

VT100 terminal



GNOME terminal in Ubuntu



# Getting help

## Man pages

```
1 ~$ man man
```

```
MAN(1)                                Manual pager utils                               MAN(1)

NAME
    man - an interface to the system reference manuals

SYNOPSIS
    man [man options] [[section] page ...] ...
    man -K [apropos options] regexp ...
    man -K [man options] [section] term ...
    man -f [whatinfo options] page ...
    man -l [man options] file ...
    man -wl-W [man-options] page ...

DESCRIPTION
    man is the system's manual pager.  Each page argument given to man is normally the name of a program, utility or function.  The manual page associated with each of these arguments is then found and displayed.  A section, if provided, will direct man to look only in that section of the manual.  The default action is to search in all of the available sections following a pre-defined order (see DEFAULTS), and to show only the first page found, even if page exists in several sections.
```

## GNU Info

```
1 ~$ info info
```

```
Next: Stand-alone Info, Up: (dir)
```

```
Stand-alone GNU Info
*****
```

```
This documentation describes the stand-alone Info reader which you can use to read Info documentation.
```

```
If you are new to the Info reader, then you can get started by typing 'H' for a list of basic key bindings. You can read through the rest of this manual by typing <SPC> and <DEL> (or <Space> and <Backspace>) to move forwards and backwards in it.
```

```
* Menu:
```

```
* Stand-alone Info::
```

```
What is Info?
```

```
* Invoking Info::
```

```
Options you can pass on the command li
```

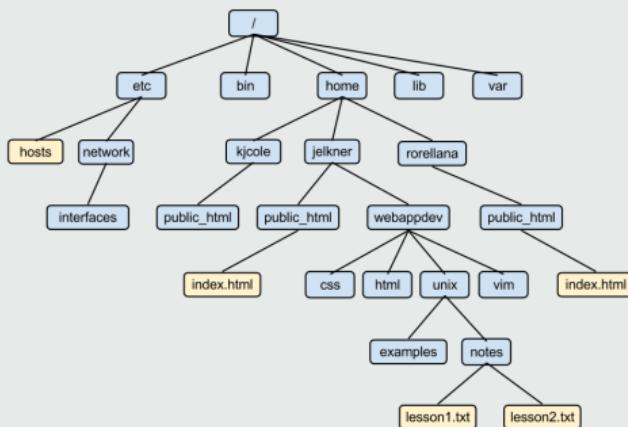
```
ne.
```

```
* Cursor Commands::
```

```
Commands which move the cursor within \
```

# Navigating the directories

## Typical directory structure



List the contents

```
1 ~$ ls [OPTION]... [FILE]...
```

Move to a directory

```
1 ~$ cd [PATH]
```

Show directory tree

```
1 ~$ tree [PATH]
```



# Users and permissions

```
> ls --color=auto -l
total 1664
drwxr-xr-x 2 wsmitt staff 4096 Apr 12 2020 000_dapi2/
-rw-r--r-- 1 wsmitt staff 16 Aug 10 2020 DisplayData
-rwxr-xr-x 1 wsmitt staff 1141 Oct 19 2019 ETG2influx.pl*
-rw-r--r-- 1 wsmitt staff 7247 Apr 15 19:22 ETGmonitor.html
-rwxr-xr-x 1 wsmitt staff 7980 Apr 9 16:11 a.out*
drwxr-xr-x 3 wsmitt staff 4096 Apr 12 23:15 bin/
-rw-r--r-- 1 wsmitt staff 76 Apr 15 19:20 birdingFromLog.sh
lrwxrwxrwx 1 wsmitt staff 12 Apr 15 12:02 curr.log -> log/0129
-rw-r--r-- 1 wsmitt staff 10333 Apr 15 19:24 dialoger.html
-rw-r----- 1 wsmitt staff 1853 Sep 6 2020 dead.letter
-rw-r--r-- 1 wsmitt staff 3975 Apr 15 16:04 dircolors.in
drwxrwxrwx 2 wsmitt staff 4096 Apr 15 12:23 drop here/
drwxr-xr-x 2 wsmitt staff 4096 Jan 29 2020 etc/
-rw-r--r-- 1 wsmitt staff 41 Jun 4 2020 index.html
-rw-r--r-- 1 wsmitt staff 11078 Mar 26 23:16 index.html.1
-rw-r--r-- 1 wsmitt staff 9749 Mar 26 23:15 'install?downld=amd.tar.gz'
-rw-r--r-- 1 wsmitt staff 6940 Jun 16 2020 locals5.06161600.gz
drwxr-xr-x 3 wsmitt staff 4096 Jun 28 2020 log/
-rw-r--r-- 1 wsmitt staff 7980 Apr 9 16:11 robin.mp3
-rw-r--r-- 1 wsmitt staff 394 Apr 9 16:11 s.c
-rw-r--r-T 2 wsmitt staff 2637 Jul 11 2020 sdr
-rw-r--r-- 1 wsmitt staff 693 Apr 15 12:05 swampRuubi.dat.gz
-rw-r--r-- 1 wsmitt staff 2637 Apr 15 19:20 swampmonitor.html
-rwxr-xr-x 1 wsmitt staff 796 Jun 1 2020 swampu.pl*
-rwsr-xr-x 1 wsmitt staff 11076 Aug 7 2020 updateCriteria*
```

Change your password

```
1 | ~$ passwd
```

See your username

```
1 | ~$ whoami
```

Change permissions

```
1 | ~$ chmod [OPTION]... MODE...
FILE...
```

read	write	execute
r	w	x
4	2	1

# Working with files

Create (empty) file

```
1 | ~$ touch FILE
```

Remove file

```
1 | ~$ rm FILE
```

Move or rename file

```
1 | ~$ mv SOURCE DEST
```

Copy file

```
1 | ~$ cp SOURCE DEST
```

Concatenate and print

```
1 | ~$ cat [FILE]
```

View file

```
1 | ~$ less FILE
```



# Interacting with remote systems

## SSH

You can log into a shell on a remote computer using secure shell (ssh)

```
1 ssh [remote user name]@[remote host name or IP address]
```

## Copying files between machines

ssh can be used to move files between computers with scp

```
1 # from a remote machine
2 scp [remote user name]@[remote host]:[remote PATH] [local PATH]
3 # to a remote machine
4 scp [local PATH] [remote user name]@[remote host]:[remote PATH]
```

# File editors

## Vim



1 | ~\$ vimtutor

## Nano

```
GNU nano 2.1.2-svn      File: ./Download/SVN/nano/src/nano.c

/* Disable mouse support. */
void disable_mouse_support(void)
{
    mousemask(0, NULL);
    mouseinterval(oldinterval);
}

/* Enable mouse support. */
void enable_mouse_support(void)
{
    mousemask(ALL_MOUSE_EVENTS, NULL);
    oldinterval = mouseinterval(50);
}

/* Initialize mouse support. Enable it if the USE_MOUSE flag is set,
 * and disable it otherwise. */
void mouse_init(void)
{
    if (ISSET(USE_MOUSE))

^G Get Help ^O WriteOut ^R Read File ^Y Prev Page ^K Cut Text ^C Cur Pos
^X Exit    ^J Justify   ^W Where Is ^V Next Page ^U Uncut Text ^T To Spell
```

# Pipes and IO redirection

Instead of printing the results of a command you can feed them to another command

*count how many files or directories are in the current directory*

```
1 | ~$ ls -l | wc -l
```

Most commands can take STDIN and output both STDOUT and STDERR  
*redirect STDOUT or STDERR to a file*

```
1 | ~$ ls 1> stdout
```

```
1 | ~$ ls [FOLDER] 2> stdout
```

*redirect BOTH and forget about it*

```
1 | ~$ ls [FOLDER] &> /dev/null
```

# GREP

Search for **regular expression** in file

How many times the word "we" is mentioned in the Federalist papers?

```
1 | ~$
```



# GREP

Search for **regular expression** in file

How many times the word "we" is mentioned in the Federalist papers?

```
1 | ~$ grep -i "we" federalist.txt
```

Pattern not specific enough

```
well aware that it would be disingenuous to resolve indiscriminately the  
jealousies and fears. So numerous indeed and so powerful are the causes which  
serve to give a false bias to the judgment, that we, upon many occasions, see  
wise and good men on the wrong as well as on the right side of questions of the
```



# GREP

Search for **regular expression** in file

How many times the word "we" is mentioned in the Federalist papers?

```
1 | ~$ grep -i "\swe\s" federalist.txt
```

But now we missed some

```
And yet, however just these sentiments will be allowed to be, we
we ought to seek it in a division of the States into distinct
```



# GREP

Search for **regular expression** in file

How many times the word “we” is mentioned in the Federalist papers?

```
1 | ~$ grep -i -E "^\w+e\s|\s\w+e\s|\s\w+e$" federalist.txt | wc -l
```

Look up the man for wc to check what we did



# sed

Lets uppercase all occurences of "we"

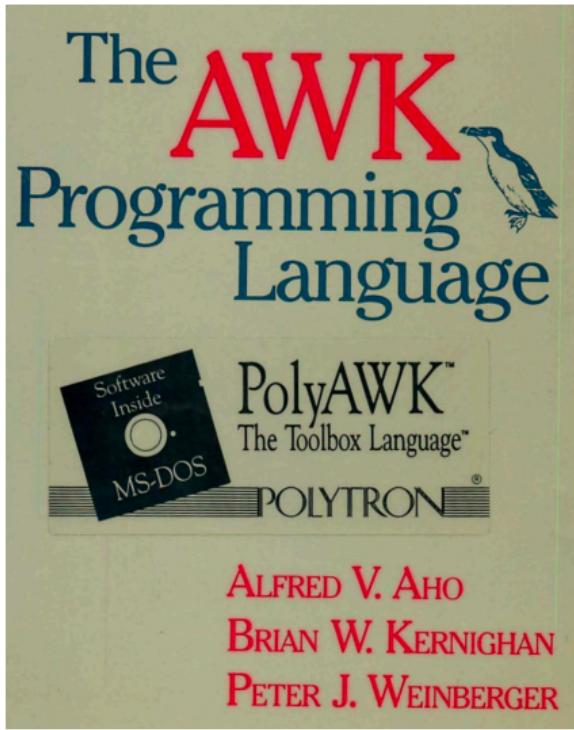
```
1 ~$ sed -e "s/\swe\s/_WE/g" federalist.txt
2 ~$ sed -e "s/^we\s/_WE/g" federalist.txt
3 ~$ sed -e "s/\swe$/_WE/g" federalist.txt
```

Pipes can stich these together

```
1 ~$ sed -e "s/\swe\s/_WE/g" federalist.txt | sed "s/^we\s/_WE/g" |
    sed "s/\swe$/_WE/g"
```



# AWK



AWK is a domain-specific language designed for text processing and typically used as a data extraction and reporting tool. Like sed and grep, it is a filter, and is a standard feature of most Unix-like operating systems.

```
1 ~$ awk 'program' FILE...
```

# AWK

## Lets parse a VCF file

```
##fileformat=VCFv4.3
##fileDate=05122018_15h52m43s
##source=IGSRpipeline
##reference=ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh38_reference_genome/GRCh38_full_analysis_set_plus_decoy_hla.fa
##INFO=<ID=AF,Number=A,Type=Float,Description="Estimated allele frequency in the range (0,1)">
##INFO=<ID=AC,Number=A,Type=Integer,Description="Total number of alternate alleles in called genotypes">
##INFO=<ID=NS,Number=1,Type=Integer,Description="Number of samples with data">
##INFO=<ID=AN,Number=1,Type=Integer,Description="Total number of alleles in called genotypes">
##INFO=<ID=EAS_AF,Number=A,Type=Float,Description="Allele frequency in the EAS populations calculated from AC and AN, in the range (0,1)">
chr1 864083 . T C . PASS AF=0.97;AC=4957;NS=2548;AN=5096;EAS_AF=1.0;EUR_AF=1.0;AFR_AF=0.99;AMR_AF=1.0;VT=SNP;DP=2591
chr1 964609 . A C . PASS AF=0.88;AC=4563;NS=2548;AN=5096;EAS_AF=0.92;EUR_AF=0.94;AFR_AF=0.78;AMR_AF=0.91;SAS_AF=0.92;VT=SNP;DP=
chr1 1065194 . G A . PASS AF=0.71;AC=3598;NS=2548;AN=5096;EAS_AF=0.78;EUR_AF=0.72;AFR_AF=0.64;AMR_AF=0.78;SAS_AF=0.66;VT=SNP;DP=
chr1 1163334 . C G . PASS AF=0.43;AC=2173;NS=2548;AN=5096;EAS_AF=0.3;EUR_AF=0.59;AFR_AF=0.28;AMR_AF=0.65;SAS_AF=0.43;VT=SNP;DP=1
chr1 1250144 . T C . PASS AF=0.09;AC=445;NS=2548;AN=5096;EAS_AF=0.17;EUR_AF=0.07;AFR_AF=0.06;AMR_AF=0.05;SAS_AF=0.09;VT=SNP;DP=2
chr1 1366262 . A G . PASS AF=0.02;AC=95;NS=2548;AN=5096;EAS_AF=0.0;EUR_AF=0.05;AFR_AF=0.0;AMR_AF=0.0;SAS_AF=0.04;VT=SNP;DP=35526
```

# AWK

Filter out header with grep

```
1 | ~$
```

Parse out the column with allele frequencies using awk

```
1 | ~$
```

Parse out the field **AF** from corresponding column

```
1 | ~$
```

# AWK

Filter out header with grep

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | less
```

Parse out the column with allele frequencies using awk

```
1 | ~$
```

Parse out the field **AF** from corresponding column

```
1 | ~$
```

# AWK

Filter out header with grep

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | less
```

Parse out the column with allele frequencies using awk

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | awk '{print $8}'  
| less
```

Parse out the field **AF** from corresponding column

```
1 | ~$
```

# AWK

Filter out header with grep

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | less
```

Parse out the column with allele frequencies using awk

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | awk '{print $8}'  
| less
```

Parse out the field **AF** from corresponding column

```
1 | ~$ grep -v "^#" common_variants_grch38_small.vcf | awk '{print $8}'  
| awk -F ";" 'print $1' | less
```