

Cursul #01

Utilizarea sistemului de fișiere



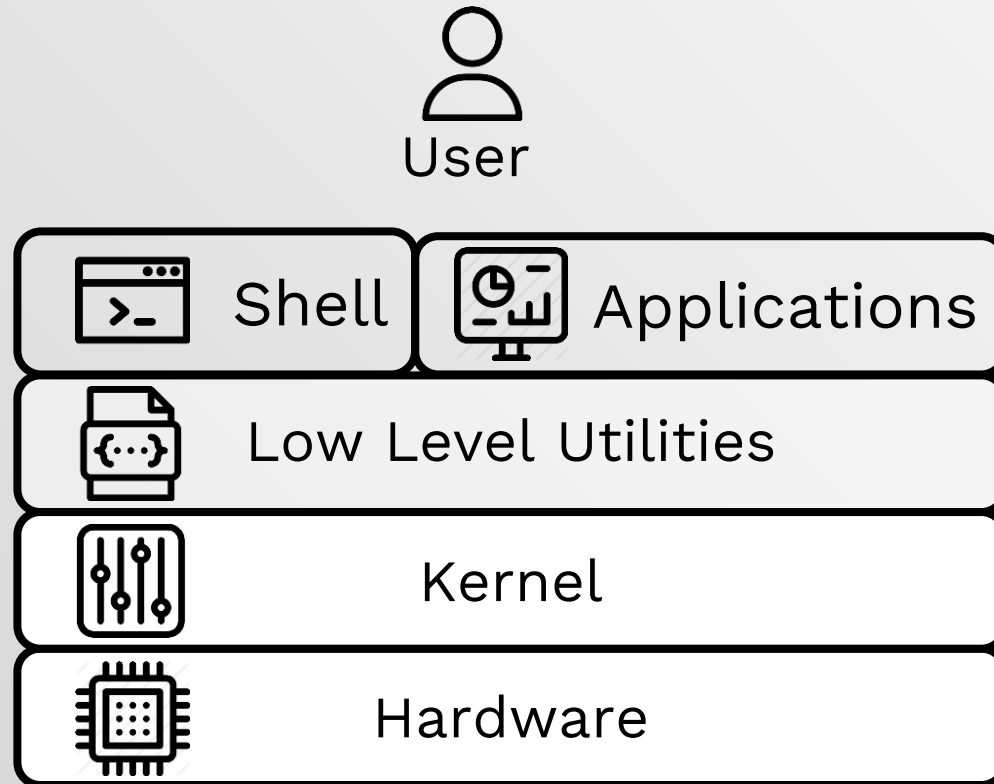
Best file compression around:

"DEL *.*" = 100% compression

Suport de curs

- Utilizarea sistemelor de operare
 - Capitolul 2 – Utilizarea sistemului de fișiere

Recapitulare



Recapitulare

Windows

Step 1. Reboot

Did that fix it?

No? Proceed to step 2

Step 2.

Format hard drive.

Reinstall Windows.

Lose all your files. Quietly weep.

Apple

Step 1. Take it to an Apple store.

Did that fix it?

No? Proceed to step 2

Step 2. Buy a new Mac.

Overdraw your account. Quietly weep.



Linux

Step 1.

Learn to code in C++. Recompile the kernel. Build your own microprocessor out of spare silicon you had lying around. Recompile the kernel again. Switch distros. Recompile the kernel again but this time using a CPU powered by refracted light from Saturn. Grow a giant beard. Blame Sun Microsystems. Turn your bedroom into a server closet and spend ten years falling asleep to the sound of whirring fans. Switch distros again. Abandon all hygiene. Write a regular expression that would make other programmers cry blood. Learn to code in Java. Recompile the kernel again (but this time while wearing your lucky socks).

Did that fix it?

No? Proceed to step 2

Step 2.

Revert back to using Windows or a Mac.

Quietly weep.

De ce?

- Sistemele de fișiere sunt banale, dar...
- Critice pentru funcționarea sistemului de operare
 - Permit controlul unei **cantități** tot mai mari de date
 - **Separă resursele** între utilizatorii multipli: umani și non-umani

Ce este un fișier?



Ce este un fișier?

- Un document digital
 - O poză, o melodie, un clip, o scrisoare, o pagină dintr-un jurnal de călătorie
 - Mai multe **tipuri**
- O formă de organizare digitală a datelor
 - O separare a informațiilor în **fragmente cu sens** pentru utilizator și **gestionabile** de sistem
 - Au forma unei înșirui de **octeți**

De ce folosim fișiere?

- Pentru a stoca informația în mod **persistent**
 - Astfel aceasta nu se pierde la reboot
- Pentru a controla informația
 - Informațiile folosite direct de **utilizatorii umani**:
poze, documente etc
 - Informațiile folosite de **aplicații**
 - Fișiere de configurare, baze de date, sunete

Ce este un director?

- O **colecție** de fișiere și subdirectoare
 - Identificată printr-un nume
 - En.: *folder, directory*
- Permit organizarea **ierarhică** a informației
 - Organizare ierarhică vs. plată
- Dacă fișier = **foaie**, atunci director = **dosar**
 - Directoarele pot fi goale
 - Directoarele pot conține alte directoare
 - **Dar** directoarele sunt tot fișiere în Linux și Mac OS X

Sistemul de fișiere

- Controlează totalitatea directoarelor și fișierelor
 - Numirea lor
 - Atributele lor
 - Accesul
- Oferă o metodă de organizare **fizică** și **logică** a fișierelor într-un mediu de stocare
- Are o organizare **ierarhică**:
 - Director rădăcină (root directory)
 - Subdirectoare
 - Fișiere

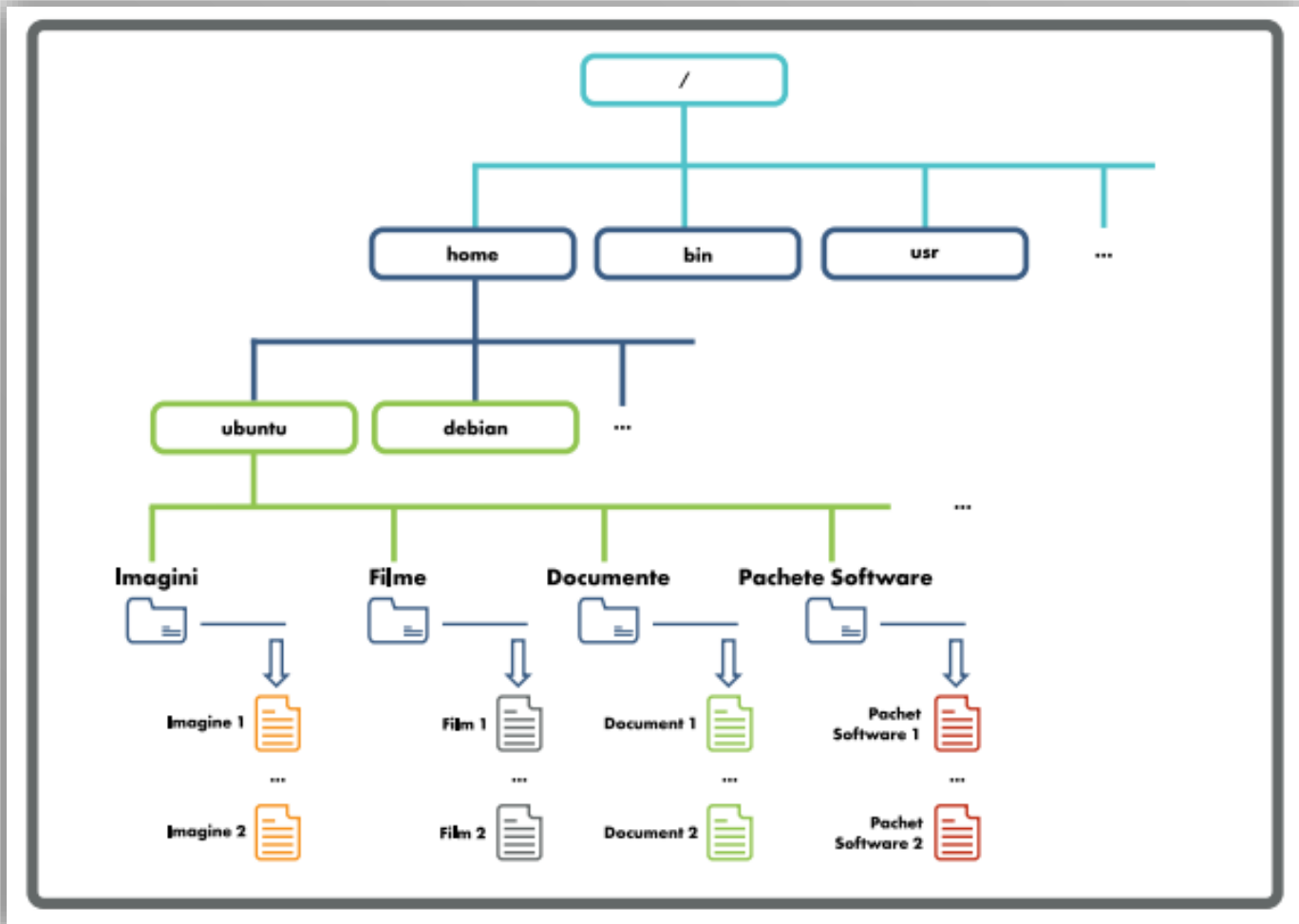
Structură plată vs ierahică



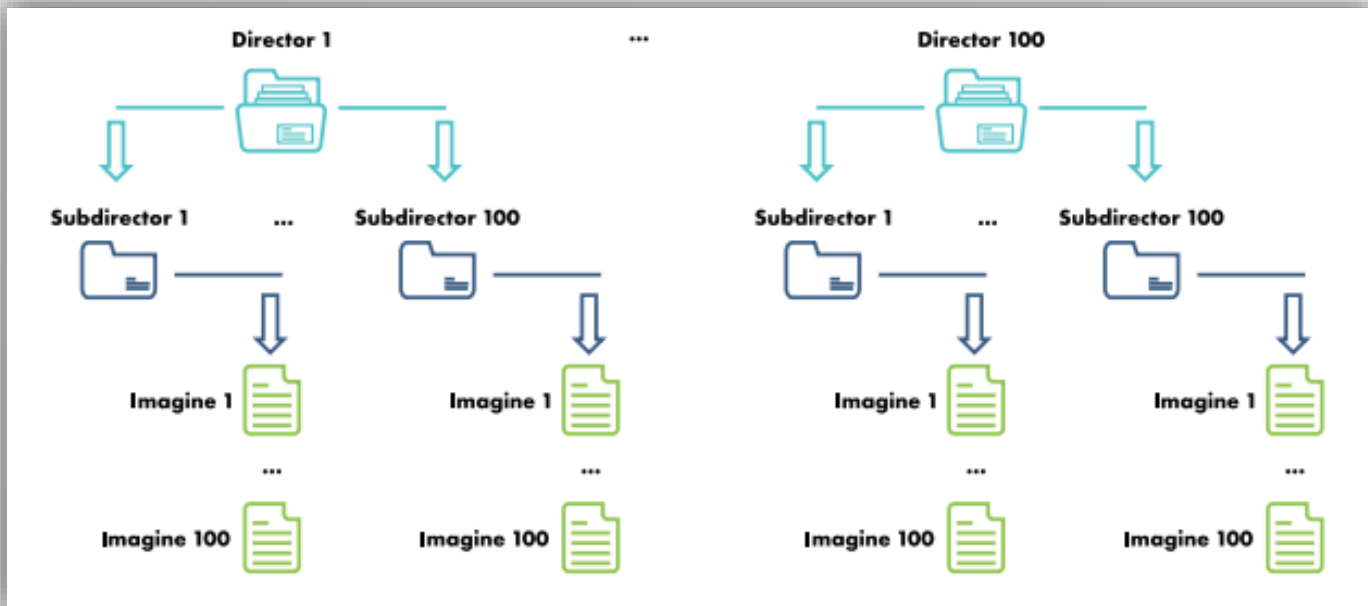
Structura plată



Structura ierahică



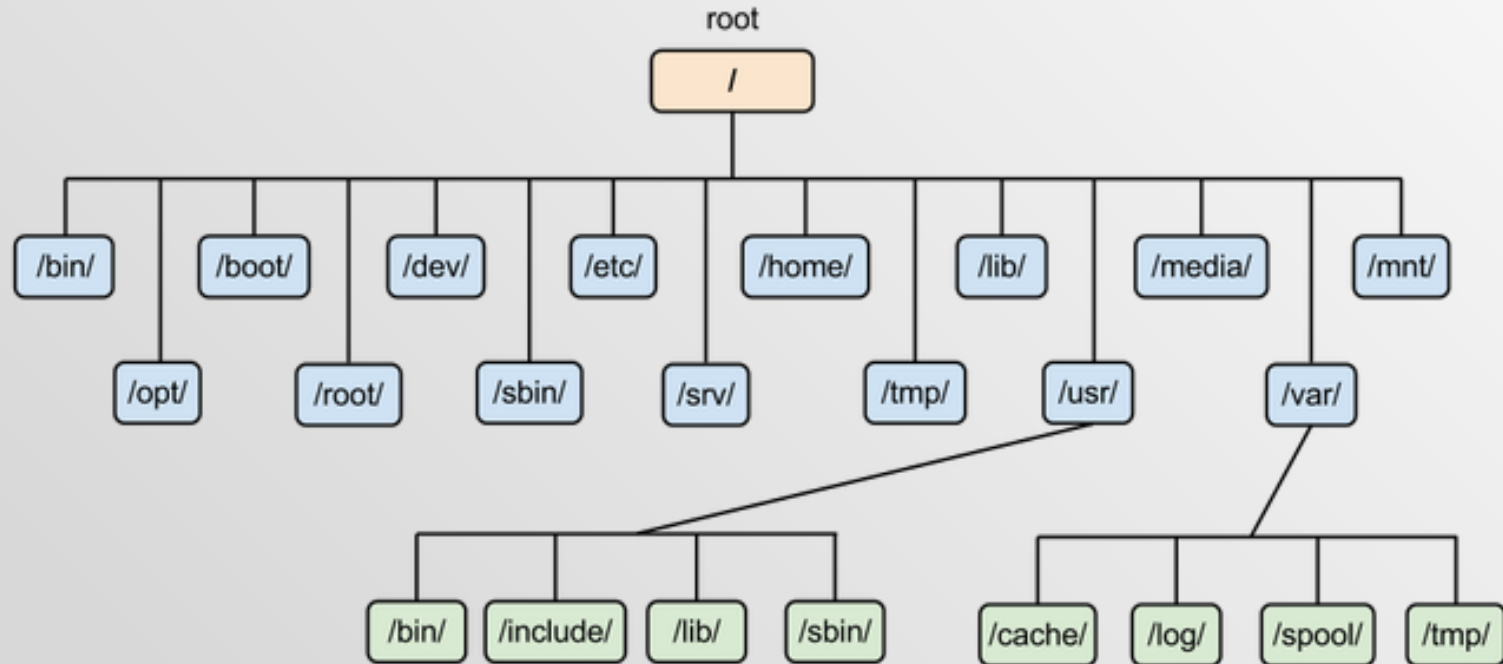
Exemplu: 1.000.000 poze



Ce ne dorim de la sistemul de fișiere?

- Să mergă repede, să ajungem repede unde ne dorim
- Să putem găsi ușor informația
- Să putem face operații multiple (creare, ștergere, redenumire etc)
- Să ofere securitate (să nu ștergem fișiere aiurea) și partiționare (să avem spații de lucru separate pentru utilizatori)

Ierarhia sistemului de fișiere (Unix/Linux)



Ierarhia în Windows

Director	Conținut
C:\	directorul rădăcină
C:\Windows	Windows-ul și fișierele aferente
C:\Documents and Settings	configurările utilizatorilor și date specifice acestora
C:\Program Files	aplicații
C:\Windows\System32	drivele și fișiere de configurare Windows
C:\Documents and Settings\username\My Documents	datele unui utilizator (aceasta este calea implicită, ea poate fi modificată)

Directorul rădăcină

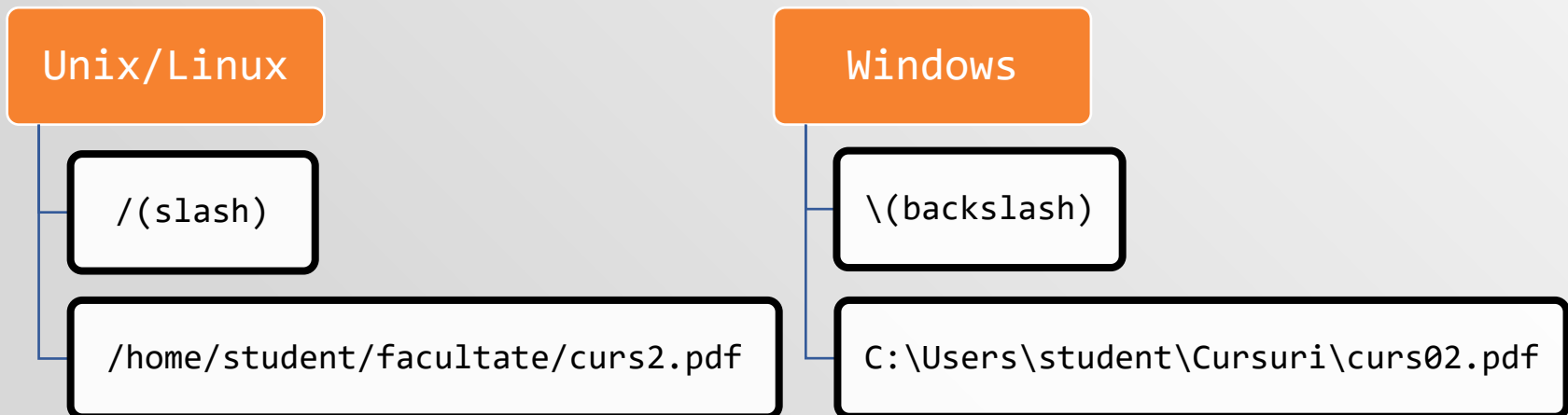
- **Linux și Mac OS:** un singur director rădăcină
- **Windows:** câte o rădăcină pentru fiecare partiție
 - A, B: floppy
 - C: hard disk
 - D sau alte litere după hard disk-uri: CD-ROM, DVD-ROM
 - E: USB drive

Directorul home

- Fiecare utilizator are un director home
- /root este pentru root
- /home/student este pentru student ș.a.m.d
- Este locul în care un utilizator obișnuit poate face ce dorește
- „~” este o scurtătură pentru directorul home al utilizatorului curent

Cale în sistemul de fișiere

- **path**
- o succesiune de intrări în sistemul de fișiere despărțite prin separator
- intrările intermediare sunt directoare, iar cele finale pot fi directoare sau fișiere



Tipuri de căi

- Absolute
 - Pleacă de la directorul rădăcină: / sau ~
 - /home/student/facultate/an-1/uso/lab-01-notite.txt
 - când e cale depărtată și ajungem mai repede pornind de la rădăcină
- Relative
 - Pleacă de la directorul curent
 - ../uso/lab-01-notite.txt
 - utile când știm cum este poziționată destinația față de directorul curent și e apropiat

Intrările . și ..

- "." și ".." sunt două directoare speciale ce se găsesc în fiecare director
- "." = indică spre același director
- ".." = indică spre directorul părinte

Navigare în sistemul de fișiere

- pwd: print working directory
 - Unde mă aflu?
- ls: list directory
 - Unde pot merge?
- cd: change directory
 - Cum ajung acolo?

```
student@uso:~$ ls
facultate
student@uso:~$ cd facultate/
student@uso:~/facultate$ ls
an-1 an-2 an-3 an-4
student@uso:~/facultate$ cd an-1/uso/
student@uso:~/facultate/an-1/uso$ pwd
/home/student/facultate/an-1/uso
```


Formatul fișierelor

- Pentru utilizator: poze, texte, melodii etc
- Pentru sistem: fișiere **text** / fișiere **binare**
 - Fișiere text
 - Text simplu (plain text): alb.txt
 - Cod sursă: alb.c, alb.java
 - Formate de prezentare: alb.html
 - Fișiere binare = non-text
 - Imagini, programe executabile, fișiere comprimate etc
- De regulă **formatul** e asociat cu **extensia**

Date binare vs Date text



Date Text

- sample.c și notes.txt
- conținutul poate fi vizualizat de om
- conține caractere tipăribile
- create de editoare de text (Vim, Notepad++, Gedit)



Date Binare

- scan.jpg și curs-02.pdf
- conținutul nu are sens pentru om
- conține caractere de orice fel (octeți, bytes)
- în general mai compacte
- create de utilitare specifice (suită Office, compilatoare, editoare de imagini)

```
student@uso:~/facultate/an-1/pc$ file sample-printf.c
sample-printf.c: C source, ASCII text
```

Extensia fișierelor

- Extensii: .txt, .tex, .mp3, .bmp, .png etc.
 - Windows: .exe = executabil
 - În Linux executabilele nu au extensie
- Interfață grafică alege programul de deschidere în funcție de extensie

Schimbarea tipului

- Dacă schimbăm manual extensia nu înseamnă că schimbăm tipul
- Putem schimba tipul deschizând cu un program și salvând ca alt tip
 - .bmp în .png
 - .docx în .pdf
 - .mp3 în .rar (comprimare, numită și arhivare)
 - .pdf în .rtf (OCR)
 - .pdf în .mp3 (text to voice)

Atribute pentru un fișier

Fișierul	uso-curs-01.pdf
Nume	uso-curs-01.pdf
Extensie	.pdf
Conținut	-
Dimensiune	1.9M
Tip de fișier	image, arhivă, binar, text, etc (format pdf, binar)
Informații de securitate	Permisuni de acces Cine a creat / modificat fișierul Unde au fost editate ultima dată

Toate cele de mai sus, mai puțin conținutul, se cheamă **metadata**. Conținutul se cheamă **date**.

„We kill people based on metadata”



General Michael Hayden, former Director of the National Security Agency and Director of the Central Intelligence Agency, 2014

<https://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/>

Director vs fișier

- aceleași metadata
- directoarele nu au, în general, extensie
- directoarele conțin alte fișiere și directoare (nu conțin date)

Dimensiunea datelor

- datele sunt stocate/transferate în biți sau octeți
- bit: valoare de 0 sau 1
- byte (octet): 8 biți, între 0 și 255 (11111111)

2^{10} octeți = 1024 octeți = 1KB (kilobyte)

2^{20} octeți = 1048576 octeți = 1MB (megabyte)

Gigabyte, terrabyte, petabyte, exabyte

Operații uzuale cu fișiere



Operații uzuale cu fișiere

- afișarea conținutului unui fișier - `cat`
- deschiderea unui fișier într-o aplicație
- editarea unui fișier – `vim`, `gedit`, etc
- crearea unui fișier - `touch`
- ștergerea unui fișier - `rm`
- copierea unui fișier - `cp`
- mutarea/Redenumirea unui fișier - `mv`

Operații pe directoare

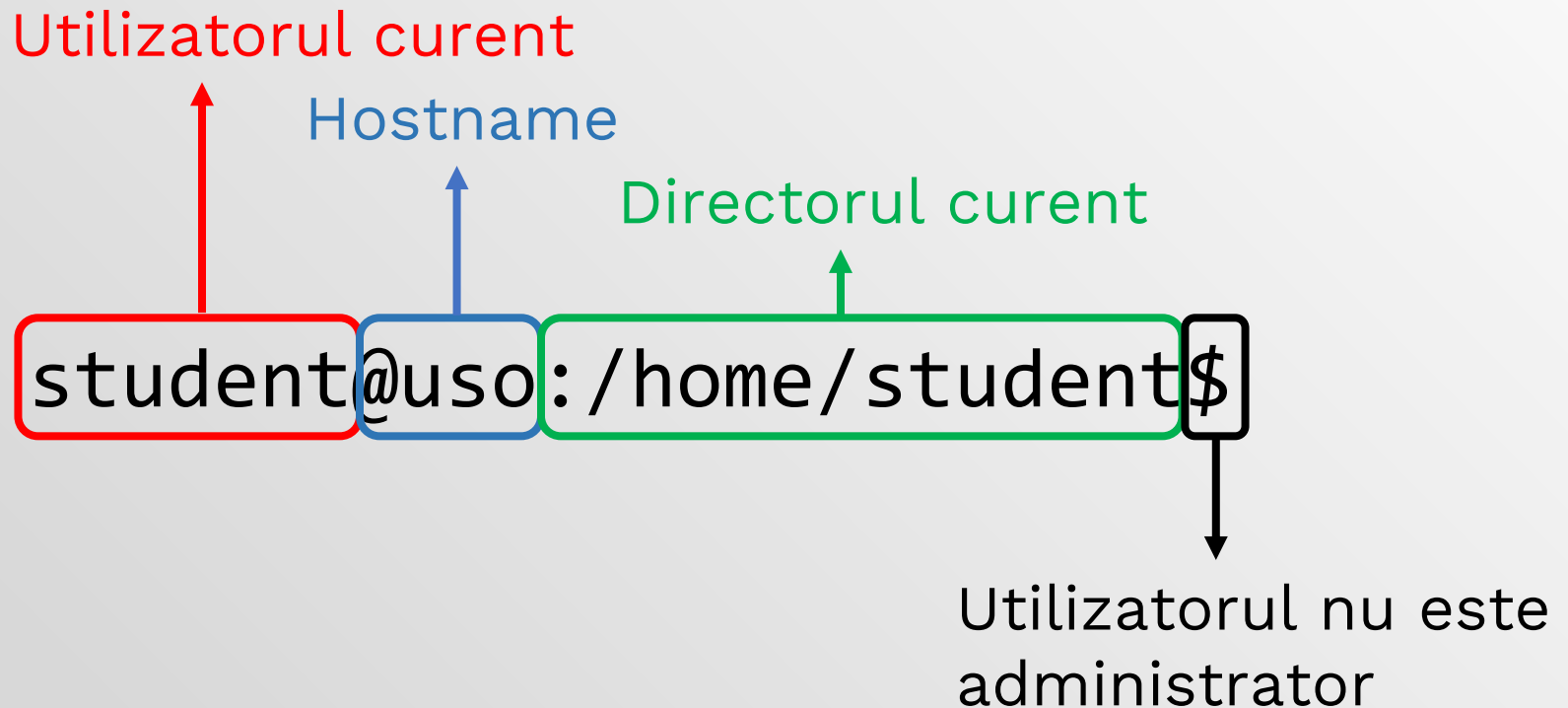
- listarea conținutului unui director - `ls`
- afișarea directorului curent - `pwd`
- schimbarea unui director - `cd`
- crearea unui director - `mkdir`
- ștergerea unui director - `rmdir`, `rm`
- copierea unui director - `cp`
- mutarea/redenumirea unui director - `mv`

Linia de comandă



- Command line interface (CLI)
- Interfață de lucru direct cu sistemul
- Alternativă la interfața grafică
- Universală
- Mai rapidă de folosit

Promptul Unix/Linux



Ce este o comandă?

- șir de caractere corespunzător unei acțiuni de executat
- poate avea parametri
- comanda poate eșua, caz în care urmărim informațiile afișate
- după orice comandă de modificare/actualizare/configurare, e indicată folosirea unei comenzi de verificare/validare (cd cu pwd, touch cu ls etc)

Tipuri de utilizatori

Utilizatori neprivilegiați

- au acces la un director home, fără alte privilegii
- au acces la directorul /tmp
- promptul se încheie în „\$” (dollar)

Utilizatorul privilegiat

- are acces la întreg sistemul de fișiere
- controlează sistemul: gestiunea pachetelor software, gestiunea de utilizatori, gestiunea rețelei
- promptul se încheie în „#” (diez, hash)

Advanced Programming in the Unix Environment

- 1992
- Richard W. Stevens (cărți depre Unix și TCP/IP)
- Steven Rago – ediția a 2-a, 2005
- Una dintre cele mai renumite cărți de programare în Unix
- Carte de căpătâi pentru programarea de sistem
- OSNews: „one of the best tech books ever published”

Theodore Ts'o

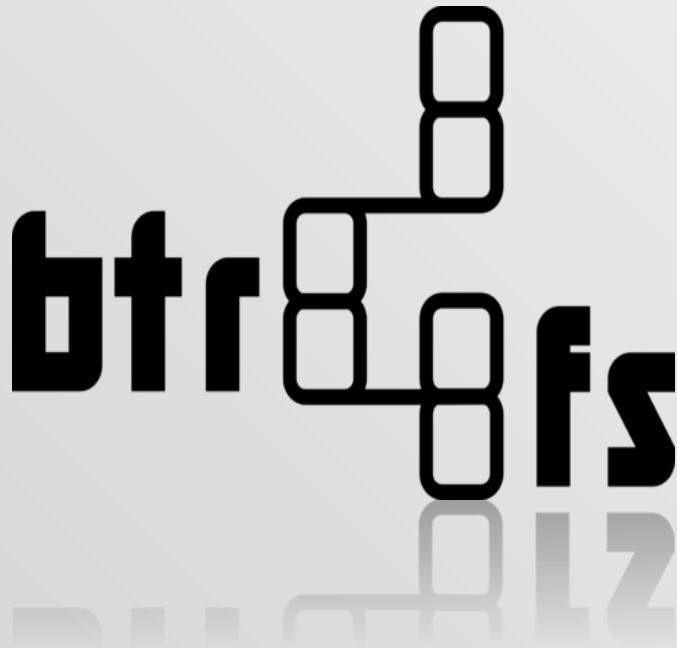
- Unul dintre primii contribuatori în nucleul Linux
- Dezvoltator și maintainer al ext2, ext3, ext4
- Contribuții în zona sistemelor de fișiere
- Lucrează la google
- Autorul `/dev/random/`



The Oracle logo, featuring the word "ORACLE" in a bold, red, sans-serif font with a registered trademark symbol (®) at the end.

ORACLE®

- www.oracle.com
- Larry Ellison
- Înființat în 1977
- Al treilea producător de software (după Microsoft și IBM)
- „flagship product” – Oracle Database
- A cumpărat Sun Microsystems în ianuarie 2010: servere, Java, MySQL, ZFS



- B-tree file system
- Copy-on-write file system
- Open source
- Început de Oracle în 2007
- Încă instabil
- Facilități similare cu ZFS de la Sun:
- Snapshots
- Partiții foarte mari
- Creșterea/Reducerea dinamică a partițiilor
- Viitorul sistem de fișiere pe Linux

Cuvinte cheie

- fișier
- director
- prompt
- comandă
- nume
- extensie
- metadate
- date
- stat
- date binare
- date text
- ascii
- bit, byte
- file
- operații
- cat
- ls
- touch
- mkdir
- cp
- mv
- rm
- rmdir
- pwd
- cd
- ierarhie
- director rădăcină
- cale relativă și absolută
- separator
- . și ..