# Notes LINFO1252

## Quentin Bodart

# Q1 2024-2025

# Contents

1	СМ	1	3
	1.1	Le syst	tème informatique et le rôle du système d'exploitation
		1.1.1	Fondamentaux
		1.1.2	Architecture de von Neumann
		1.1.3	Fonctionnement d'un système informatique 4
		1.1.4	Traitement d'une interruption
		1.1.5	Accès direct à la mémoire
		1.1.6	Système informatique complet
		1.1.7	Rôle du système d'exploitation
		1.1.8	Virtualisation
		1.1.9	Séparation entre mécanisme et politique 5
		1.1.10	Modes d'exécution
		1.1.11	Appel système
	1.2		tion de la ligne de commande
		1.2.1	Utilitaires UNIX
		1.2.2	Shell / Interpréteur de commandes
		1.2.3	Flux et redirections
		1.2.4	Scripts
2	CM	2 : lang	guage C et compléments 10
	2.1		guage C
		2.1.1	Préprocesseur
		2.1.2	headers
		2.1.3	Intégration avec le shell
		2.1.4	Affichage formatté
	2.2	Types	de données
		2.2.1	Nombres signés et flottants
		2.2.2	Caractères
		2.2.3	Chaînes de caractères
		2.2.4	Pointeurs
		2.2.5	Pointeurs vers une fonction
		2.2.6	Manipulation de bits

3	CM3 : Gestion de la mémoire				
	3.1	Organ	isation d'un programme Linux en mémoire	15	
	3.2 Gestion de la mémoire dynamique		n de la mémoire dynamique	16	
		3.2.1	Objectif d'un algorithme de gestion de mémoire dynamique .	16	
		3.2.2	Implémentation de l'algorithme	16	

# Objectifs du cours

- utiliser et comprendre les systèmes informatiques (i.p. GNU/Linux)
- utiliser les services fournis par les SE (systèmes d'exploitation)
- design et mise en oeuvre d'un SE

## 1 CM 1

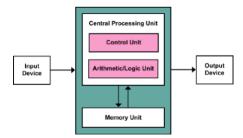
## 1.1 Le système informatique et le rôle du système d'exploitation

**Syllabus**: https://sites.uclouvain.be/SystInfo/notes/Theorie/intro.html

#### 1.1.1 Fondamentaux

- Composants :
  - CPU / Processeur
  - Mémoire Principale (RAM)
  - Dispositifs d'entrée/sortie (y.c. de stockage)
- Fonctionnement d'un CPU
  - Lire / écrire en mémoire vers / depuis des registres
  - Opérations (calculs, comparaisons) sur ces registres
- Jeux d'instructions
  - x86\_64 (PC, anciens Mac)
  - ARM A64 (Raspberry PI, iPhone, nouveaux Mac M1-3)

#### 1.1.2 Architecture de von Neumann

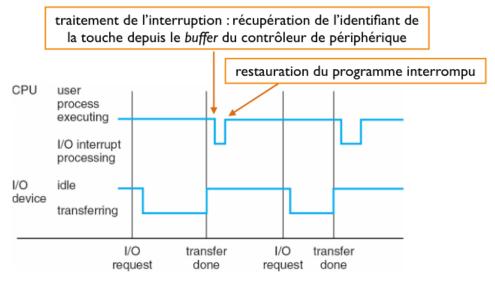


#### 1.1.3 Fonctionnement d'un système informatique

- La représentation des données se fait en binaire.
- Les opérations d'entrée-sortie se déroule de manière concurrente.
- Il y existe des contrôleurs distinct controlant chacun un type de périphérique.
- Chaque contrôleur possède une mémoire dédiée (buffer)
- Le processeur doit déplacer des donnés depuis/vers la mémoire principale depuis/vers ces buffers dédiés
- Le processeur suit un "fil" continu d'instructions
- Le contrôleur de périphérique annonce au processeur la fin d'une opération d'entrée/sortie en générant une interruption (signal électrique à destination du processeur)

#### 1.1.4 Traitement d'une interruption

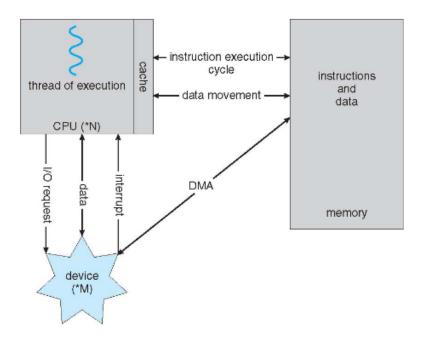
Le processeur interrompt le fil d'exécution d'instructions courant et transfert le contrôle du processeur à une routine de traitement. Cette même routine détermine la source de l'interruption, puis restaure l'état du processeur et reprend le processus :



#### 1.1.5 Accès direct à la mémoire

Direct Memory Access (DMA) désigne le fait de ne pas faire une interruption à chaque octet lu depuis un disque dur. Une interruption est tout de même faite à la fin du transfert d'un **bloc**.

#### 1.1.6 Système informatique complet



#### 1.1.7 Rôle du système d'exploitation

Programmer directement au-dessus du matériel, gérer les interruption, etc... serait une trop grosse tâche pour le programmeur.

#### 3 rôles principaux

- Rendre l'utilisation et le développement d'applications plus simple et plus universel (portable d'une machine à une autre)
- Permettre une utilisation plus efficace des ressources
- Assurer l'intégrité des données et des programmes entre eux (e.g., un programme crash mais pas le système)

#### 1.1.8 Virtualisation

Le système d'expoitation **virtualise** les ressources matérielles afin de fonctionner de la même manière sur des systèmes avec des ressources et composants fort différents. Chaque SE doit trouver un compromis entre abstraction et efficacité!

#### 1.1.9 Séparation entre mécanisme et politique

• Un mécanisme permet le partage de temps

 Une politique arbitre entre les processus pouvant s'exécuter et le(s) processeur(s) disponibles

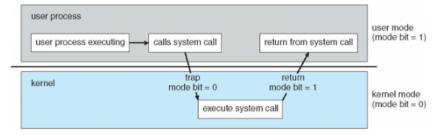
On peut définir des politiques d'ordonnancement différentes selon les contextes, mais sur la base du même mécanisme.

#### 1.1.10 Modes d'exécution

- mode utilisateur : programme utilisant les abstractions fournies par le SE ; certaines instructions sont interdites, comme par exemple:
  - Accès à une zone mémoire non-autorisée (SegFault)
  - De manière générale, toutes les instructions permettant de changer la configuration matérielle du système, comme la configuration ou la désactivation des interruptions
- mode protégé : utilisé par le noyau du SE, toutes les instructions sont autorisées
- L'utilisation de fonctionnalités du SE par un processus utilisateur nécessite de passer d'un mode à l'autre : **appel système**

#### 1.1.11 Appel système

- Un appel système permet à un processus utilisateur d'invoquer une fonctionnalité du SE
- Le processeur interrompt le processus, passe en mode protégé, et branche vers le point d'entrée unique du noyau :



## 1.2 Utilisation de la ligne de commande

 $\textbf{Syllabus}: \ \texttt{https://sites.uclouvain.be/SystInfo/notes/} \ \ \textbf{Theorie/shell.html}$ 

#### 1.2.1 Utilitaires UNIX

La philosophie lors de la création des utilitaires UNIX était de créer des outils les plus simples possible, donc d'avoir une seule tâche par outil :

# **Quelques utilitaires standard**

Utilitaire	Fonction	
cat	lire/afficher le contenu d'un fichier ex:cat fichier.txt	
echo	afficher une chaîne de caractères passée en argument, ex.:echo "Bonjour Monde"	
head / tail	affiche le début resp. la fin d'un fichier ex. : tail errors.log	
wc	compte le nombre de caractères / de lignes d'un fichier. ex.:wc -l students.dat	
sort	trie un fichier. ex.:sort -n -r scores.txt	
uniq	extrait les lignes uniques ou dupliquées d'un fichier <b>trié</b> fourni en argument. ex. : uniq -d students.dat	

Afin d'en savoir plus sur

unecommande, il suffit d'utiliser l'utilitaire 'man'.

#### 1.2.2 Shell / Interpréteur de commandes

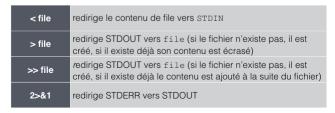
Rend possible l'interaction avec le SE. Il en existe plusieurs, mais le principal est 'bash'. Il est toujours complémentaire à une **interface graphique**.

#### 1.2.3 Flux et redirections

## Flux standards et redirections



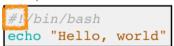
- 3 flux standards (I entrée, 2 sorties)
- Redirections permettent de combiner des commandes en utilisant des fichiers intermédiaires



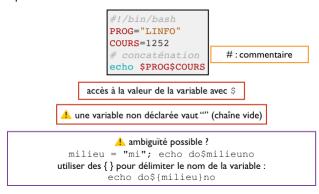
Exemple de redirections :

#### 1.2.4 Scripts

Un système UNIX peut exécuter du language machine ou des **languages interprétés** Un script commence par convention par les symboles #!, référant à l'interpréteur, ici bin/bash :



Ils peuvent aussi contenir des variables:



et des conditionnelles :

## **Conditionnelles (if)**



#### et des **boucles for**:

```
#!/bin/bash
# exemple_for.sh
students="Jule Maxime Hakim"
for s in $students; do
    l=`wc -l TPl-$s.txt | cut -d' ' -fl`
    echo "Bonjour $s, ton compte rendu de TP comporte $l lignes."
done
```

- s prend successivement les valeurs présentes dans la liste d'entrée \$students
- `command` permet d'assigner la sortie d'une commande à une variable
  - que fait la commande composée ci-dessus ?
- boucles while et until: principe similaire, avec une condition d'arrêt booléenne

## 2 CM2 : language C et compléments

## 2.1 Le language C

Jusqu'en 1970, les OS étaient codés en assembleur, ce qui était très chronophage et difficile à maintenir, d'où la nécessité d'un language plus haut niveau, mais toujours proche de la machine : le C. Son objectif était de concevoir un language :

- Haut niveau (structuré)
- Proche du matériel (accès direct à la mémoire et traduction direct au processeur)
- Efficace (non-interprété et sans machine virtuelle)
- Portable

#### 2.1.1 Préprocesseur

Le code C est tout d'abord transformé par un préprocesseur avant d'être compilé en langage machine par le compilateur.

Les directives telles que "#include" ou "#define" y sont remplacées par du texte spécifique à la commande.

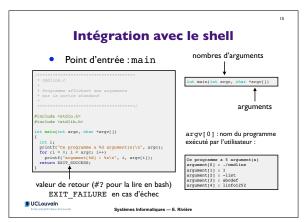
C'est seulement après ça que le code est compilé en code machine.

#### 2.1.2 headers

Les directives #include du préprocesseur permettent d'importer des définitions de fonctions, de constantes, etc. d'une librairie Séparation entre fichiers headers (.h) et sources (.c) :

- Le header contient la défnition des fonctions, suffisante pour savoir comment générer le code pour les appeler
- Le code source contient la mise en œuvre de ces fonctions
- #include "monheader.h" pour inclure un header du dossier courant (du même projet)
- #include <header.h> pour inclure un header standard

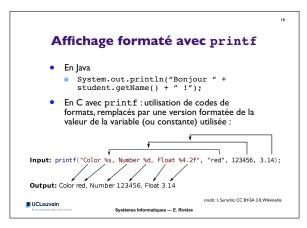
## 2.1.3 Intégration avec le shell



Le tableau d'argument donné à main se termine toujours par NULL.

## 2.1.4 Affichage formatté

Pour afficher un String formatté en C, on peut utiliser la fonction printf:



Modificateurs (format specifiers):

%с	char
%d / %i	signed int
%e / %E	scientific notation
%f	float
%ld / %li	long
%lf	double
%Lf	long double

## 2.2 Types de données

• int, long : variable entière

• char : charactère

• double, float : nombre en virgule flottante

• booléens : définis dans stdbool.h, true ou false

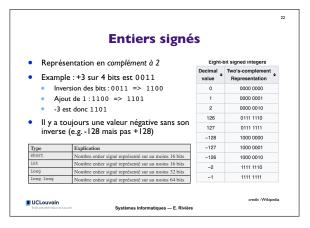
ullet String : pas d'objet String en c, seulement une chaîne de char o longueur de la chaîne inconnue !

• Structures de contrôle : while, for, if, ...

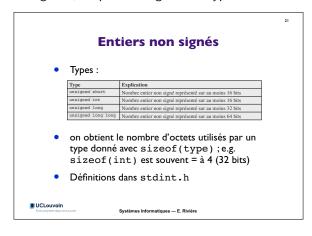
La taille occupée par un type de donnée peut être récupérée via sizeof().

## 2.2.1 Nombres signés et flottants

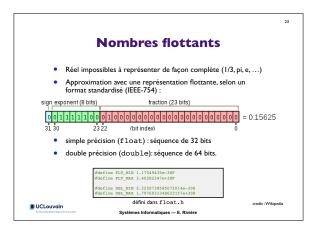
Les entiers, en C, sont signés par défaut :



Via le mot-clé "unsigned", on peu "dé-signer" un type de donnée :



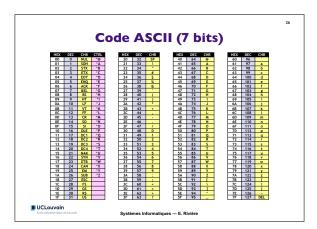
Représentation de nombres flottants:



## 2.2.2 Caractères

Les caractères en C sont représentés par un nombre, en accord avec la norme ASCII. Deux types de charactères :

- 0x0.....: permet de représenter les caractères anglais
- 0x1.....: permet de représenter les caractères accentués



#### 2.2.3 Chaînes de caractères

En C, une chaîne de charactères est stockée sous la forme d'un tableau. Le dernier élément du tableau contient le charactère " $\$ 0".

#### 2.2.4 Pointeurs

Un **pointeur** est une variable qui contient l'adresse à laquelle est stockée un autre élément (variable, fonction, etc.)

La mémoire peut être vue comme une suite de "cases" (octets) avec des adresses consécutives (index de chaque octet).

- On peut créer un pointeur en utilisant le caractère "\*" et récupérer une adresse en utilsant le caractère "&".
- Le charctère "\*" peut aussi être utilisée sur un pointeur pour récupérer sa valeur (déréférencement)
- En pratique, char[] et char\* sont équivalents.
- Les pointeurs permettent de passer une variable par référence à une fonction.

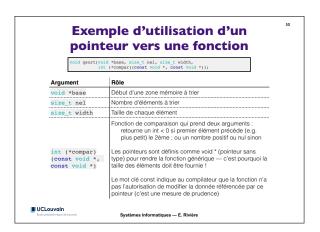
#### 2.2.5 Pointeurs vers une fonction

Une fonction contient une séquence d'instructions, commençant à l'adresse où est stockée sa première instruction.

On peut stocker l'adresse d'une fonction dans un **pointeur de fonction**:

 ${\sf D\'eclaration: type (*ptr)([type\_arg]+)}$ 

Exemple:



#### 2.2.6 Manipulation de bits

4 opérateurs binaires :

• tilde : inversion  $\rightarrow$  tilde00000000 = 111111111

• & : ET  $\rightarrow$  11111010 & 01011111 = 01011010

 $\bullet \ | : \ \mathsf{OU} \to \mathsf{11111010} \ | \ \mathsf{01011111} = \mathsf{11111111}$ 

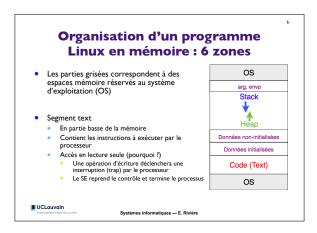
• : OU exclusif  $\rightarrow$  11111010 01011111 = 10100101

## 3 CM3 : Gestion de la mémoire

## 3.1 Organisation d'un programme Linux en mémoire

Les différentes parties de la mémoire de Linux sont appelés **Segments**. Ceux qui nous intéressent:

- Segment 1 : code
   Il contient les instructions à exécuter par le processeur, directement extraites des codes.
- Segment 2 : données initialisées
   Contient les données initialisées par les codes du segment 1
- Segment 3 : données non-initialisées
   Contient les données non-initialisées par le codes du segment 1.
   Elles sont souvent initialisées à 0 par défaut, en attente d'être transférées dans le segment 2.
- Segment 4 : Heap et Stack
  - Heap : Permet aux programmes d'y réserver des zones mémoire (c.f. malloc, free, valgrind)
  - Stack : Permet de stocker les variables locales et les appels de fonctions, ainsi que de récupérer les valeurs de retour.
- Segment 5 : Arguments et variables d'environnement



## 3.2 Gestion de la mémoire dynamique

malloc() et free() ne sont pas mises en oeuvre dans le noyau du système d'exploitation, ce sont simplement des fonctions da la librairie standard utilisant des appels système : brk() et sbrk().

malloc() est aligné sur un **facteur d'alignement** : elle retourne un nombre entier d'octets directement supérieur au nombre d'octets demandés. (padding) Exemple : Sous Linux, ce facteur est de **16 octets**. Si l'on demande 37 octets, le malloc va en retourner 64.

#### 3.2.1 Objectif d'un algorithme de gestion de mémoire dynamique

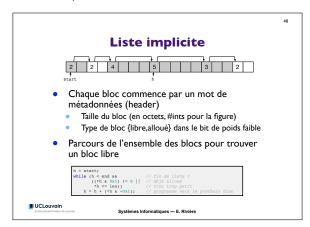
Les objectifs de cet algorithme sont les suivants :

- Minimiser le temps d'exécution et maximiser sa stabilité
- **Utiliser efficacement la mémoire** (minimiser la fragmentation): Deux types de fragmentation:
  - Fragmentation interne : espace perdu par le padding
  - Fragmentation externe : espace perdu par l'allocation et désallocation de la mémoire
- Bonnes propriétés de localité spatiale :

Alloue les variables proches les unes des autres afin d'accélérer l'accès du cache du processeur

#### 3.2.2 Implémentation de l'algorithme

Une des implémentations pourrait utiliser des métadonnées :



Cependant, aucune des implémentations ne peut satisfaire les 3 conditions parfaitement. Par exemple :

