

DIU – EIL

Version du 08 juillet 2019

Maël Guiraud – Mael.Guiraud@uvsq.fr
Franck Quessette – Franck.Quessette@uvsq.fr

Tout le contenu de ce document suppose de travailler sous linux avec une fenêtre shell.

1 Partie 1 – Fichier

Les caractéristiques principales d'un fichier sont :

- son **inode** : numéro unique dans le système ;
- son nom ;
- son type : fichier regular, directory, pipe, link, ... ;
- son propriétaire et son groupe d'appartenance ;
- ses droits.

On ne rappelle pas ici les commandes de base (création, effacement, recopie, déplacement) sur les fichiers et les répertoires.

Exercice 1 *Caractéristiques des fichiers*

Utiliser la commande `ls` avec l'option `-l` pour visualiser ces informations dans le répertoire courant.

Exercice 2 *Modification des droits*

- Utiliser la commande `touch` suivi d'un nom de fichier (disons `toto2019`) pour créer un fichier vide.
- Utiliser la commande `ls` avec l'option `-l` pour visualiser les droits du fichier `toto2019`.
- Utiliser la commande `chmod 000 toto2019` pour modifier les droits du fichier.
- De nouveau `ls -l`.
- Essayer d'effacer le fichier `toto2019`.
- Utiliser la commande `chmod 541 toto2019` pour modifier les droits du fichier.
- De nouveau `ls -l`.

Exercice 3 *Écriture dans un fichier en C*

Taper ce programme dans un éditeur puis le compiler et l'exécuter.

```
#include <stdio.h>
#include <stdlib.h>
int main () {
    FILE *F;
    F = fopen("toto28","w");
    if (!F) exit(1);
    fprintf(F,"Bonjour\n");
    fclose(F);
    exit(0);
}
```

Utiliser `ls -l` pour voir les droits du fichier `toto28`

2 Partie 2 – Processus

Les caractéristiques principales d'un processus sont :

- son **pid** : process ident, numéro unique dans le système, attribué en ordre croissant ;
- son processus parent ;
- son état : **R** Running, **S** Sleeping, **Z** Zombi, ... ;
- son propriétaire.

Exercice 4 Caractéristiques des processus

- Utiliser la commande **ps -u** pour visualiser certains processus.
- Utiliser la commande **ps -u** pour visualiser de nouveaux les processus, quelle différence.
- Utiliser la commande **ps -ejf** pour visualiser plus de processus. Repérer les **pid** et les **ppid** des processus **ps** et **bash**.
- Utiliser la commande **ps -ejf | grep user** pour filter les lignes.
- Ajouter la commande **wc** pour compter le nombre de processus du système d'une part et le nombre de processus de **user** d'autre part.

Exercice 5 Création de processus en C

Dans le programme ci-dessous, repérer le type : **pid_t** et les fonctions : **fork()**, **getpid()**, **getppid()**, **wait()**.

```
#include <stdio.h>
#include <stdlib.h>
#include <sys/types.h>
#include <sys/wait.h>
#include <unistd.h>

int main() {
    pid_t mon_pid;
    int i;
    printf("Dans pere avant = %d %d\n",getpid(),getppid());

    mon_pid = fork();
    printf("mon_pid = %d\n",mon_pid);
    if (mon_pid==0) {
        printf("Dans fils = %d %d\n",getpid(),getppid());
        for (i=0 ; i<10 ; i++) {
            printf("fils\n");
            sleep(1);
        }
    }
    else {
        printf("Dans pere = %d %d\n",getpid(),getppid());
        for (i=0 ; i<5 ; i++) {
            printf("pere\n");
            sleep(2);
        }
        printf("Pere : debut attente de fin du fils\n");
        wait(NULL);
        printf("Pere : fin attente de fin du fils\n");
    }
}
```

```
}  
  
    exit (0);  
}
```

- Utiliser la commande `man` pour comprendre ce que font chacune de ces fonctions.
- Comment utiliser `getpid()` pour mettre un verrou sur un fichier ou une application.
- Comment utiliser `getpid()` pour initialiser un générateur pseudo-aléatoire.

3 Partie 3 – Interface web en ligne de commande

Exercice 6 Premier `wget`

- Créer un dossier `WEB2019`.
- Aller dans ce dossier.
- Taper la commande `wget http://www.uvsq.fr`. Que s'est-il passé ?
- Ouvrir le fichier `index.html` avec un navigateur. Que constatez-vous ?
- Quelle différence entre du php et du html ?
- Qu'est ce qui est transmis par le réseau ?
- Qui interprète le php ?

Exercice 7 Protocole IP, DNS, Adresse IP

- Taper la commande `hostname -i`.
- Taper la commande `wget -O monip1.html ifconfig.me` (le `http://` est facultatif).
- Taper la commande `wget -O monip2.html checkip.dyndns.org`.
- A qui sert le DNS ?
- Qu'utilisent généralement les fournisseurs d'accès en France pour empêcher les internautes d'aller sur certains sites (de torrent par exemple) ?
- Comment changer son DNS ? Sous linux `/etc/resolv.conf` ou avec `nmcli`.

4 Partie 4 – Mini projet

Le but est d'écrire un script shell pour récupérer plusieurs pages web du site `https://www.okkazeo.com` afin d'afficher les jeux vendus dans une ville donnée par son code postal. Essayer de respecter ces étapes et de bien les comprendre avant de passer à la suivante

Étape 1 : Récupérer une page

1. Aller sur le site `https://www.okkazeo.com` pour se familiariser avec celui-ci.
2. Sur le site cliquer sur "Autour de moi" puis sélectionner un département. Constater le changement dans l'URL.
3. Aller sur une autre page que la page 1 du département choisi et constater le changement dans l'URL.

À cette étape vous devriez être capable à la main de récupérer toutes les pages d'un département donné

Étape 2 : Comprendre le contenu d'une page

1. Avec `wget` récupérer la page 2 du département 78 dans un fichier `78.2.html`.
2. Ouvrir la page récupérée avec un éditeur et essayer de comprendre comment est structurée la page et où sont les infos sur les jeux.

Étape 3 : Modifier le contenu de la page avec la commande `grep`

La commande `grep` prend en argument une chaîne de caractères et un nom de fichier et affiche les **lignes** du fichier contenant la chaîne de caractères. Afin de mettre le résultat dans un nouveau fichier, il faut rediriger la commande :

1. Taper la commande `grep "Disponible" 72.2.html > 78.2.dispo.html`.
2. Que contient le fichier `78.2.dispo.html` ? regarder avec l'éditeur et avec le navigateur

Étape 4 : Filtrer par code postal la commande `sed`

Les commandes `grep` et `sed` travaillent sur les lignes. On va donc modifier le fichier pour mettre un jeu par ligne.

1. Taper la commande `sed -e 's/<article/\n<article/' 78.2.dispo.html > 78.2.jeu.html`

La commande `sed` avec l'option `-e` effectue l'action d'édition donnée entre apostrophe. Ici l'action est la substitution d'où le `s`. La syntaxe est `sed -e 's/chaîne1/chaîne2/' fic1 > fic2`. Cette commande remplace sur chaque ligne la première occurrence de `chaîne1` par `chaîne2`. On peut forcer `sed` à remplacer toutes les occurrences et non pas la première en ajoutant un modificateur d'action à la fin. Ici on va rajouter un `g` pour signifier global :

1. Taper la commande `sed -e 's/<article/\n<article/g' 78.2.dispo.html > 78.2.jeu.html`

Étape 5 : Filter les lignes sur un code postal avec la commande `grep`

1. Utiliser la commande `grep` pour ne garder que les lignes contenant le code postal `78140`.

Étape 6 : Enchaîner les actions dans un script shell

1. Créer un fichier `filtre.sh` contenant les lignes suivantes :

```
#!/bin/bash

codepostal="$1"
echo "Code Postal = $codepostal"
departement=$((codepostal/1000))
echo "Departement = $departement"

page=2
url="https://www.okkazeo.com/jeux/arrivages?departement=$departement&page=$cpt"
```

```
echo "Url = $url"
wget -q -O 78.html "$url"
sleep(1)
```

2. Comprendre le script (utiliser man pour les options de wget).
3. Rendre le fichier exécutable.
4. Exécuter le script : `./filtre.sh 78140`
Remarque importante : Quand on scripte la récupération de pages web, il vaut mieux mettre une temporisation entre deux wget afin de ne pas saturer le serveur avec une fausse manip.
5. Ajouter les autres commandes pour filter le résultat.

Étape 7 : Boucle en bash

1. Tester et comprendre ce script (rechercher “comparaison bash” dans google pour comprendre les comparaisons) :

```
#!/bin/bash

cpt=1
while [ "$cpt" -le "10" ]; do
    echo "cpt = $cpt"
    cpt=$((cpt+1))
done
```

2. Rendre le fichier exécutable.
3. Utiliser la boucle dans votre script `filtre.sh` pour récupérer toutes les pages d’un département.
4. Comment s’arrêter automatiquement sur le bon nombre de pages ?

Étape 8 : Finalisation et plus

1. Ajouter des filtres sur le résultat produit pour que l’affichage soit plus propre et plus correct.
2. Récupérer toutes les pages d’un département et créer un fichier par code postal existant.
3. Dans les filtre garder la date de publication des annonces.
4. Aller rechercher sur le site boardgamegeek les notes des jeux. et les afficher.