

6 - Textutils

Laurent Tichit

5 avril 2011

Plan

1 Fichiers texte

Fichiers texte

Affichage, Concaténation

- **cat fich1** : Affiche le fichier sur la sortie standard.
- **cat fich1 fich2 fich3...** : Concatène plusieurs fichiers.

Début et fin de fichier

- **head [-n] [fichier]** : affiche les n premières (10 par défaut) lignes du fichier.
- **tail [[+ nombre] [lbc]] [fichier]** : recopie sur stdout le morceau de fichier spécifié.

Affichage page par page

- **more [-n] fichier...** : permet d'afficher page par page un long fichier.
- **less fichier...** : permet d'afficher page par page un long fichier.
Permet les mêmes commande interactives que **man**.

Tri de lignes

- **sort [options] [[+n1] [-n2]] [fich1...]** : permet de trier lexicographiquement les lignes des fichiers d'entrée.
tichit@iml230:~\$ **sort -t: +2n /etc/passwd**
 - ▶ Trie le fichier /etc/passwd
 - ▶ sur le troisième champ (+2n),
 - ▶ séparateur ':' (-t:)
- **uniq [options] fichier1 fichier2** : permet de supprimer les lignes *adjacentes* identiques d'un fichier.

Découpage par colonne

- La commande **cut** permet d'extraire certains champs d'un fichier.
 - ▶ **-c** suivant le nombre de caractères
 - ▶ **-f** suivant le nombre de champs
 - ▶ **-d** spécifie le séparateur de champ
- **cut -c-10 /etc/passwd**
Extrait les 10 premiers caractères de chaque ligne.
- **cut -c2-10 /etc/passwd**
Extrait les caractères 2 à 10 de chaque ligne.
 - ▶ **n** : n^{ieme} caractère compté à partir de 1.
 - ▶ **m-** : du m^{ieme} caractère jusqu'à la fin de la ligne.
 - ▶ **m-n** : du m^{ieme} au n^{ieme} (inclus) caractère.
 - ▶ **-n** : du premier au n^{ieme} (inclus) caractère.
- **cut -d: -f 5 /etc/passwd**
Extrait le champ 5 de chaque ligne. Le séparateur de champs est ' : '.

Concaténation verticale

- **paste fich1 fich2...** : Concatène ligne à ligne plusieurs fichiers. En fait chaque fichier devient la colonne d'un tableau.

```
tichit@iml230:~$ cat premier
```

```
1
```

```
1
```

```
1
```

```
tichit@iml230:~$ cat deuxieme
```

```
2
```

```
2
```

```
2
```

```
tichit@iml230:~$ paste premier deuxieme
```

```
1 2
```

```
1 2
```

```
1 2
```

Comptage des caractères, mots, lignes

- **wc [options] [fichier...]** : compte le nombre de lignes, de mots ou de caractères d'un fichier.

Comparaison

- **diff [options] fichier1 fichier2** : permet de comparer deux fichiers ligne à ligne.

Commande **tr**

Recopier l'entrée sur la sortie en remplaçant certains caractères par d'autres.

- Remplace les 'a' par des 'A', les 'b' par des 'B' et les 'c' par des 'C' :
tr abc ABC < fichier1
- Remplace les majuscules par des minuscules :
tr "[A-Z]" "[a-z]" < fichier1
- Placer chaque mot d'un fichier donné sur une ligne différente.
tr -cs "[A-Za-z-']" "\n" < fichier1
- Créer un dictionnaire de tous les mots d'un texte en ordre alphabétique.
tr -cs "[A-Za-z-']" "[\n]" | sort | uniq