Commande Awk

Cette commande agit comme un filtre programmable, prenant une série de lignes en entrée (par fichier, ou via l'entrée standard) et écrivant sur la sortie standard (redirigeable). Awk prend en entrée les lignes une par une et choisit les lignes à traiter (ou non) par des expressions rationnelles ou des numéros de lignes. Une fois la ligne sélectionnée, elle est découpée en champs, selon un séparateur d'entrée indiqué dans le programme awk par le symbole FS (par défaut espace et tabulation). Puis les différents champs sont disponibles dans les variables ($1, $2, $3, $NF pour le dernier). Vous pourrez aussi trouver des fichiers .awk qui sont des scripts écrits dans ce langage.

```
awk [-F] '{action-awk}' [ fichier1 fichier2 ..... fichiern ]
```

La commande prend en paramètre la liste des fichiers à traiter, si des fichiers ne sont pas spécifiés, awk travaillera sur l'entrée standard. On peut donc placé la commande derrière un tube de communication. L'option "-F" permet d'initialiser si besoin la variable "FS" (Field Separator).
Voici une liste des variables prédéfinies au lancement de awk:

Variables	Valeur par défaut	Rôle
RS	Newline (\n)	Record Separator : Caractère séparateur d'enregistrement (lignes)
FS	Suite d'espaces et ou de tabulation	Field Separator : caractères séparateurs de champs.
OFS	Espace	Output Field Separator: Séparateur de champ pour l'affichage
ORS	Newline (\b)	Output Record Separator : Séparateur de ligne pour la sortie
ARGV	-	Tableau initialisé avec les arguments de la ligne de commande (awk inclus)
ARGC	-	Nombre d'arguments dans le tableau ARGV
ENVIRON	Variables d'environnement exportées par le shell	Tableau contenant les variables d'environnement exportées par le shell.
CONVFMT	%.6g	Format de conversion des nombres en string
OFTM	%.6g	Format de sortie des nombres.
SUBSEP	\0.34	Caractère de séparation pour les routines internes de tableaux.

Attention : Lorsque que la variable FS a plus de 2 caractères, elle sera interprété comme une expression régulière
Voici les variables qui sont initialisées lors du traitement d'une ligne

Variable	Rôle
$0	Valeur de l'enregistrement courant
NF	$1 contient la valeur du 1er champ, $2 du 2ème, ... et $NF le dernier
$1,$2, .. $NF	$1 contient la valeur du 1er champ, $2 du 2ème, ... et $NF le dernier
NR	Number : indice de l'enregistrement courant (NR vaut 1 quand la 1ère ligne est en cours de traitement)
FNR	File Number : indice de l'enregistrement courant relatif au fichier en cours de traitement
FILENAME	Nom du fichier en cours de traitement
RLENGTH	Longueur en string trouvé par la fonction match()
RSTART	Première position du string trouvé par la fonction match

Exemple d'utilisation :
`ps -ef| awk '{print $1,$8}'`
avec ps -ef on a la liste des processus, en cours d'exécution (voici un extrait):

ZorNItUeJ9sy4ITR-image-1560346772728.png

et avec la commande:

lDRWmDBEZNzYucuV-image-1560346794846.png

La partie en gras représente entre quote permet d'éviter l'interprétation par le shell et indique à awk quelle fonction exécuter sur le traitement du awk. Les instructions doivent être placées entre accolades. La fonction intégrée print va afficher à l'écran les champs 1 et 8 de chaque ligne.
Voici un autre exemple :
`ps -ef | awk '{print "User : ", $1, "\tCommande : ", $8}'`

yS2M7xCiAxMtVh8G-image-1560346819530.png

Critères de sélection :
Il est possible de sélectionner les enregistrements sur lesquels l'action doit être exécutée : `awk [-F] 'critère {action-awk}' [fichier1 fichier2 ... fichiern]`
Le critère de sélection peut s'exprimer de différentes manières.

Expression régulières :
`awk -F':' '/\/bin\/false/ {print $0}' /etc/passwd`

C4WGKYDcaWjF0F69-image-1560346843983.png

Par défaut, c'est le champ correspondant au $0 qui est mis en correspondance avec l'expression régulière., mais il est possible de le spécifier avec l'opérateur de concordance "~" ou de non concordance "!~".

```
awk -F ':' '$6 ~ /^\/usr/ {print $0}' /etc/passwd
```

3vBz4fDvVT7V07ym-image-1560346868683.png

Test logiques :
Le critère peut être une expression d'opérateurs renvoyant vrai ou faux.
Par exemple, pour afficher uniquement les lignes impaires du fichier /etc/passwd
`awk -F':' 'NR%2 {print $0}' /etc/passwd`

Revision #1
Created 2019-10-31 12:08:19 UTC by Cécile
Updated 2019-10-31 12:08:30 UTC by Cécile