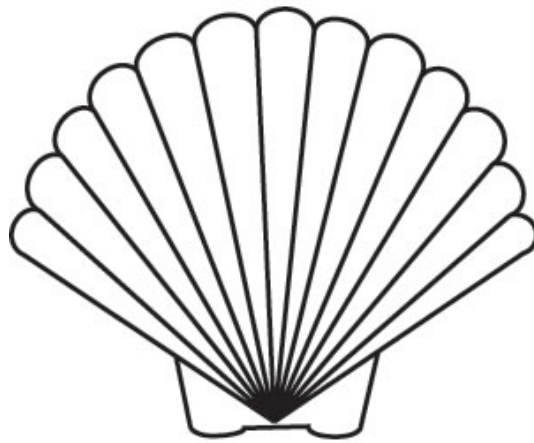


Administració de Sistemes Operatius



“Pràctica de Shell”

Vicenç Juan Tomàs Montserrat

DNI: 41518542-T

UIB - Curs 2010/11

Dr. Ricardo Galli



Índex

Introducció.....	3
Exercici 1: Tamany de cada fitxer.....	4
Exercici 2: Tràfic total per cada fitxer.....	5
Exercici 3: Tràfic total per cada IP analitzada.....	6
Exercici 4: El dia amb més tràfic.....	7
Exercici 5: L'hora amb més tràfic.....	8
Exercici 6: El dia amb més visitants (per IP diferents).....	9
Exercici 7: Fitxers sol·licitats no existents (404).....	10



Introducció

Aquesta pràctica consta de 7 exercicis que s'han de realitzar amb l'interpret Bash¹. Bash es un interpret desenvolupat per GNU² que la seva funció consisteix en interpretar ordres que li arriben de d'una consola o terminal.

Bash es un interpret amb moltes funcionalitats que permeten realitzar qualsevol funció del tipus, cercar patrons dins un fitxer, redirigir una sortida cap a un fitxer, etc.

Per a desenvolupar aquesta pràctica s'ha emprat el programa awk³ que es un llenguatge de programació dissenyat per al processament de dades basades en text.

El script o guió d'execució consta d'un menú que on es pot elegir l'exercici a executar.

Menú interactiu

- 1) Tamany de cada fitxer
- 2) Tràfic total per cada fitxer
- 3) Tràfic total per a cada direcció IP analitzada
- 4) El dia amb més tràfic (en bytes servits)
- 5) L'hora amb més tràfic (en bytes servits)
- 6) El dia amb més visitants (IP diferents)
- 7) Fitxers sol·licitats no existents (codi 404)
- Q) Sortir

Opció:

El primer que es fa es comprovar que s'ha passat per paràmetre l'arxiu correcte, es a dir, es comprova que es passa un únic paràmetre i que efectivament aquest paràmetre es un fitxer i no un directori o una imatge, per exemple. També es recull el primer paràmetre \$0 i es guarda dins la variable FITXER.

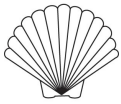
```
if [ $# -eq 0 -o ! -f "$1" ]; then
    echo "Ús: `basename $0` fitxer_log" >&2
    exit 1
fi
```

Posteriorment, ja es poden executar les funcions en que estan separades en cadascun dels exercicis.

1 BASH *Bourne-Again SHell*

2 GNU és un acrònim recursiu de l'anglès "*GNU's Not Unix*"

3 AWK deriva dels cognoms dels autors: *Alfred Aho*, *Peter Weinberger*, y *Brian Kernighan*



Exercici 1: Tamany de cada fitxer

Per a resoldre aquest exercici s'han filtrat les columnes corresponents del fitxer. En aquest cas son les columnes 7 i 10, que son la URL sol·licitada i el nombre de bytes transferits respectivament.

L'objectiu es que per a cada URL se mostri el tamany del fitxer de la URL. Per aconseguir això s'ha indexat la columna de les URL (7) i s'ha aconseguit eliminar les URL repetides de tal manera que només falta mostrar el tamany del cada fitxer indicat en la columna 10.

El codi:

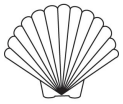
```
awk '{
    total[$7]=$10
} END {
    for (v in total)
        print "El tamany total del fitxer " v " és: " total[v], "bytes"
}' $FITXER
```

El resultat:

....

```
El tamany total del fitxer /cill/fotos/santantoni2007/simple/style.css és: 1067 bytes
El tamany total del fitxer /~gallir/CursPL/memories/ és: 338 bytes
El tamany total del fitxer /gallir/posts/2004/12/05/22/ és: 13405 bytes
El tamany total del fitxer /gallir/posts/2006/03/04/659/ és: 21309 bytes
El tamany total del fitxer /~gallir/VRML és: 326 bytes
El tamany total del fitxer /gallir/posts/2005/02/22/ és: 12144 bytes
El tamany total del fitxer /gallir/posts/2007/03/29/1022/ és: 13125 bytes
El tamany total del fitxer /gallir/posts/2005/12/20/551/ és: 9163 bytes
El tamany total del fitxer /gallir/posts/category/curiosidades/page/12/ és: 35227 bytes
El tamany total del fitxer /gallir/posts/2005/03/22/ és: 10914 bytes
El tamany total del fitxer //admin/pma/scripts/setup.php és: 338 bytes
El tamany total del fitxer /gallir/posts/2007/03/22/ és: 11069 bytes
El tamany total del fitxer /gallir/posts/2006/12/10/902/ és: 30140 bytes
El tamany total del fitxer /gallir/posts/2006/03/22/ és: 10447 bytes
```

...



Exercici 2: Tràfic total per cada fitxer

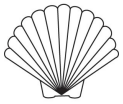
Aquest exercici es molt semblant a l'exercici anterior però en aquest cas s'ha de tenir en compte el tràfic total per cada fitxer, es a dir, que s'ha d'indexar la URL del fitxer (columna 7) i per cada URL del fitxer anar sumant els bytes transferits i retornar el total de cadascun.

El codi:

```
awk '{
    total[$7]+=$10
} END {
    for (v in total)
        print "El tràfic total del fitxer " v " és: " total[v], "bytes"
}' $FITXER
```

El resultat:

```
...
El tràfic total del fitxer /cill/fotos/santantoni2007/simple/style.css és: 22407 bytes
El tràfic total del fitxer /~gallir/CursPL/memories/ és: 4394 bytes
El tràfic total del fitxer /gallir/posts/2004/12/05/22/ és: 147455 bytes
El tràfic total del fitxer /gallir/posts/2006/03/04/659/ és: 340944 bytes
El tràfic total del fitxer /~gallir/VRML és: 3586 bytes
El tràfic total del fitxer /gallir/posts/2005/02/22/ és: 97152 bytes
El tràfic total del fitxer /gallir/posts/2007/03/29/1022/ és: 787500 bytes
El tràfic total del fitxer /gallir/posts/2005/12/20/551/ és: 256564 bytes
El tràfic total del fitxer /gallir/posts/category/curiosidades/page/12/ és: 422724 bytes
El tràfic total del fitxer /gallir/posts/2005/03/22/ és: 87312 bytes
El tràfic total del fitxer //admin/pma/scripts/setup.php és: 338 bytes
El tràfic total del fitxer /gallir/posts/2007/03/22/ és: 99621 bytes
El tràfic total del fitxer /gallir/posts/2006/12/10/902/ és: 512380 bytes
El tràfic total del fitxer /gallir/posts/2006/03/22/ és: 73129 bytes
...
```



Exercici 3: Tràfic total per cada IP analitzada

Es tracta de realitzar la mateixa funció que en el l'exercici anterior però ara envers d'indexar per la columna 7, indexar per la columna 1, la direcció IP.

El codi:

```
awk '{
    total[$1]+=$10
} END {
    for (v in total)
        print "El tràfic total de la IP " v " és: " total[v], "bytes"
}' $FITXER
```

El resultat:

```
...
El tràfic total de la IP 189.161.89.202 és: 68637 bytes
El tràfic total de la IP 85.52.253.3 és: 12576 bytes
El tràfic total de la IP 201.86.187.105 és: 48334 bytes
El tràfic total de la IP 187.171.19.212 és: 48334 bytes
El tràfic total de la IP 81.36.83.130 és: 322 bytes
El tràfic total de la IP 189.243.67.202 és: 407781 bytes
El tràfic total de la IP 189.240.77.202 és: 48334 bytes
El tràfic total de la IP 62.57.247.53 és: 34074 bytes
El tràfic total de la IP 189.222.87.202 és: 48334 bytes
El tràfic total de la IP 186.28.228.65 és: 48334 bytes
El tràfic total de la IP 88.5.26.150 és: 56936 bytes
El tràfic total de la IP 190.228.67.75 és: 48334 bytes
El tràfic total de la IP 62.43.35.192 és: 196631 bytes
El tràfic total de la IP 201.102.12.244 és: 48334 bytes
El tràfic total de la IP 201.137.6.248 és: 48334 bytes
El tràfic total de la IP 195.162.69.202 és: 1183763 bytes
El tràfic total de la IP 201.243.70.244 és: 48334 bytes
El tràfic total de la IP 62.42.16.153 és: 48334 bytes
El tràfic total de la IP 190.125.77.36 és: 57258 bytes
El tràfic total de la IP 190.203.247.71 és: 300638 bytes
...
```



Exercici 4: El dia amb més tràfic

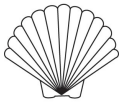
En aquest exercici s'usen les columnes 4 (data) i 10 (nombre de bytes transmesos). Per tant, s'han de filtrar aquestes dues columnes i anar sumant el tràfic per dia. Finalment calcular el màxim i treure el resultat per pantalla.

El codi:

```
awk '{
    total[$4]+=$10
} END {
    for (v in total)
        print v":"total[v]
}' $FITXER | awk -F: '{
    total[$1]+=$5
} END {
    for (v in total)
        print v,total[v]
}' | awk 'BEGIN {
    max=0
}
{
    if($2>max) {
        dia=$1; max=$2
    };
} END {
    print "El dia amb més trafic és: " dia " -> ",max" bytes "
```

El resultat:

El dia amb més trafic és: [18/Feb/2011 -> 1653471845 bytes



Exercici 5: L'hora amb més tràfic

Aquest exercici es semblant a l'anterior, però aquesta vegada es necessari trobar l'hora amb més tràfic, per tant es filtraran les columnes 4 (hora) i 10.

El procediment es el mateix que abans.

El codi:

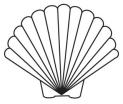
```
awk '{
    print $4,$10
}' $FITXER | awk -F: '{
    print $2,$4
}' | awk '{
    print $1,$3
}' | awk '{
    total[$1]+=$2
} END {
    for (v in total)
        print v,total[v]
}' | awk 'BEGIN {
    max=0
}
{ if($2>max) {
    hora=$1; max=$2
};
} END {
    print "La hora amb més trafic és: " hora"h ->", max" bytes "
}'
```

El resultat:

La hora amb més trafic és: 21h -> 2.54871e+09 bytes



Exercici 6: El dia amb més visitants (per IP diferents)



Exercici 7: Fitxers sol·licitats no existents (404)

En aquest cas es tracta de comptar el nombre de 404 que hi ha hagut. Per això s'ha de controlar la columna 9 que indica els resultats de les sol·licituds. No es més que realitzar un comptador per a cada coincidència amb el numero 404 i mostrar el resultat al final.

El codi:

```
awk '{
    print $9
}' $FITXER | awk 'BEGIN {
    count=0
}
{ if($1==404) {
    count+=1
}
};
END {
    print "Nombre de 404: "count
}'
```

El resultat:

Nombre de 404: 66190