

PRÁCTICA 3: Búsqueda de ficheros y patrones en *bash*

ALUMNO:

1. INTRODUCCIÓN

En *bash* tenemos tres comandos esenciales para la búsqueda de archivos y patrones dentro de archivos: *grep*, *find* y *locate*. Son parecidos entre sí pero vamos a ver las principales diferencias entre los mismos en esta práctica.

1.1. *Grep (Globally Regular Expressions Pattern)*

El comando ***grep*** busca patrones en ficheros (dentro de su contenido). Por defecto devuelve todas las líneas que contienen un patrón (cadena de texto) determinado en uno o varios ficheros. Utilizando las opciones se puede variar mucho este comportamiento. Si no se le pasa ningún fichero como argumento hace la búsqueda en su entrada estándar.

Grep [opciones] <patrón> [ficheros]

Algunas opciones son las siguientes:

Opción	Descripción
-c	Devuelve solo la cantidad de líneas que contienen al patrón
-i	Ingresa las diferencias entre mayúsculas y minúsculas
-H	Imprime, además de las líneas, el nombre del fichero donde se encontró el patrón
"^letra"	Busca todas las líneas que comienza por la letra que especifiquemos
-v	Devuelve las líneas que no contiene el patrón
-r	Busca en un directorio de forma recursiva
-n	Imprime el número de cada línea que contiene al patrón
-rn	Busca de forma recursiva en varios ficheros de archivos

1.2. *Find*

El comando *find* es uno de los más poderosos en un sistema Linux. Permite buscar de forma recursiva en un directorio a todos los ficheros que cumplan ciertas condiciones. Las condiciones pueden estar relacionadas con el nombre de los ficheros, el tamaño, los permisos, el tipo, las fechas de acceso y modificación, etc.

Algunas opciones:

1. **-name <extensión>** permite especificar patrones para los nombres de los ficheros a buscar.
2. **-iname<expresión>** permite especificar patrones para los nombres a buscar pero sin tener en cuenta mayúsculas ni minúsculas.
3. **-size +/- <n>** permite indicar el tamaño máximo y/o mínimo de los ficheros a buscar. Por defecto el tamaño se expresa en bloques de 512 bytes, pero si se precede este por un carácter *c* se referirá a bytes, *k* a kilobytes, *w* a palabras de dos bytes y *b* a bloques.
4. **-perm [+|-] <modo>** permite referirse a aquellos ficheros cuyos permisos sean exactamente modo, incluya todos los de modo (signo -) o incluya alguno de los de (signo +). El valor de se expresa en forma numérica.
5. **-exec <comando>** permite definir un comando a ejecutarse para cada resultado de la búsqueda. La cadena {} se sustituye por el nombre de los ficheros encontrados. El carácter ; permite indicar la finalización del comando. (Tanto {} como ; tienen que ir entre comillas o entre contrabarras para evitar que sea sustituido por el shell).

1.3. Locate

El comando *locate* busca en una base de datos, actualizada periódicamente, todos los *paths* en la jerarquía de ficheros que contengan una cadena determinada. En esta práctica no vamos a utilizarlo.

2. OBJETIVO

El objetivo de la práctica es que os familiaricéis con los comandos que sirven para localizar patrones dentro de archivos, así como encontrar determinados ficheros.

3. EJERCICIOS

3.1. Los primeros ejercicios los vamos a realizar con el archivo *allstar.txt* que utilizamos en las anteriores prácticas. Cuenta el número de veces que aparece la palabra “All” en el archivo.

```
grep -o 'All' allstar.txt | wc -l
```

3.2. Cuenta el número de veces que aparece la palabra “nba” en el mismo.

```
grep -o 'nba' allstar.txt | wc -l
```

3.3. Hazlo ahora sin distinguir mayúsculas de minúsculas.

```
grep -i 'All' allstar.txt | wc -l
```

3.4. Busca las líneas que no contienen la letra “n” en *allstar.txt*

```
grep -v 'n' allstar.txt
```

3.5. Imprime en pantalla todas las líneas con el número 2 de los ficheros *allstar.txt* y *fichero.txt* de la anterior práctica.

```
grep '2' allstar.txt fichero.txt
```

3.6. Busca las líneas que comienzan por la letra *a* en *allstar.txt*.

```
grep grep '^a' allstar.txt allstar.txt
```

3.7. Busca información sobre cómo podemos imprimir las líneas que comienzan por *a* pero también, a la vez, las que comienzan por *l* en *allstar.txt*.

```
grep '^a|^l' allstar.txt
```

3.8. Busca los múltiplos de 5 en el archivo *fichero.txt*.

```
grep "[05]"$ fichero.txt
```

3.9. Busca información sobre cómo se puede buscar las palabras que acaben en una determinada letra.

```
grep 'a$' archivo.txt
```

En este comando, la expresión regular *a\$* significa que estás buscando líneas que tengan la letra "a" al final de la línea.

3.10. Pon un ejemplo del anterior ejercicio para cada fichero.

```
grep 'a$' allstar.txt
```

3.11. Encuentra todos los archivos *.csv* de la carpeta *Downloads*.

```
find ~/Downloads -name "*.csv"
```

3.12. Encuentra todos los archivos de texto de la carpeta *Downloads*.

```
find ~/Downloads -name "*.txt"
```

3.13. Encontrar todos los archivos *.pdf* que tengan tamaño superior a 1MB.

```
find / -type f -name "*.pdf" -size +1M
```

3.14. Encuentra todos los archivos en formato *.pdf* o *.csv* que tengan un tamaño inferior a 500 kb. Para ello utiliza

```
find . \“-name *.pdf” -o “-name *.csv” -size -500k
```

3.15. Encuentra todas las imágenes de la carpeta.

```
find Segundo cuatri/ -type f \-iname "*.jpg"
```

3.16. Encuentra todos los ficheros que contienen la letra *x*.

```
find Segundo cuatri -type f -exec grep -l 'x' {} +
```

3.17. Encuentra todos los ficheros del directorio que sean en formato *pdf* y además tengan la letra *a*. Ejemplo: para encontrar los archivos que sean *.csv* con la letra *x* usaríamos:

```
find . -type f -name "*.pdf" -iname "*x"
```

3.18. Utiliza el siguiente comando para encontrar los archivos *.txt* que contengan más de 10 líneas:

```
find . -name "*.txt" -exec awk 'NR>10 {print FILENAME; exit}' {} \;
```

3.19. Encuentra todos los archivos *txt* modificados en los últimos 30 días añadiendo después de localizar, en general, cualquier archivo *txt* el comando *-atime -30*.

```
find -atime -30 -name "*.txt"
```

3.20. Encontrar todos los archivos PDF que hayan sido modificados en las últimas 24 horas.

```
find -atime -1 -name "*.pdf"
```