

1 Grep

O comando **grep** nos permite buscar, dentro de arquivos, as liñas que coinciden cun patrón.

Por defecto, **grep** imprime as liñas encontradas na saída estándar. É dicir, que podemos velo directamente na pantalla, ou redireccionar a saída estándar a un arquivo.

Vexamos as opcións mais empregadas:

- c Devolve só o número de liñas que coinciden, non o seu contido.
- e **PATRÓN** Permítenos especificar varios patróns de procura ou protexer aqueles patróns de procura que comezan co signo -.
- r Busca recursivamente dentro de todos os subdirectorios do directorio actual.
- v Móstranos as liñas que non coinciden co patrón buscado.
- i Ignora a distinción entre maiúsculas e minúsculas.
- l Que só mostraría os nomes dos arquivos que conteñan o obxectivo buscado.
- n Numera as liñas na saída.
- E Permítenos usar expresións regulares.
- o Indícalle a grep que nos mostre só a parte da liña que coincide co patrón.
- f **ARCHIVO** Extrae os patróns do arquivo que especifiquemos. Os patróns do arquivo deben ir un por liña.
- H Imprimémos o nome do arquivo con cada unha das coincidencias.

O programa **grep** toma a expresión que recibe como parámetro como unha **expresión regular básica**. Iso quere dicir que algúns caracteres teñen un significado especial polo que hai que andar con ollo.

Vexamos algúns exemplos:

- Buscar todas as liñas que conteñan a palabra **iface** no arquivo **/etc/network/interfaces**:

```
$ grep iface /etc/network/interfaces
# Outro xeito de facelo, sería:
$ cat /etc/network/interfaces | grep iface
```

- Buscar todas as liñas que conteñan a palabra **iface eth0** no arquivo **/etc/network/interfaces**:

```
$ grep "iface eth0" /etc/network/interfaces
```

- Buscar todas as liñas que teñan palabras que comece pola letra **i** no arquivo **/etc/network/interfaces**:

```
$ grep -E '^<i.*\>' /etc/network/interfaces
```

- Buscar todas as liñas que comece pola letra **i** no arquivo **/etc/network/interfaces**:

```
$ grep -E '^i' /etc/network/interfaces
```

- Buscar todas as liñas que rematen en por ".DS" ou ".DE":

```
$ grep -E '\.D[SE]$' arquivo
```

- Buscar todas as liñas que comece pola letra **i** no arquivo **/etc/network/interfaces**:

```
$ grep -E '^i' /etc/network/interfaces
```

- Contar o número de interfaces de rede que temos definidos no ficheiro **/etc/network/interfaces**:

```
$ grep -c "iface eth" /etc/network/interfaces
```

- Mostrar por pantalla as liñas do arquivo **/etc/apt/sources.list** que non son comentarios:

```
$ grep -v "#" /etc/apt/sources.list
```

- Buscar as liñas nas que apareza a palabra "UNIX" pero sin distinguir entre maiúsculas e minúsculas (*ignore case*):

```
$ grep -i "unix" arquivo
```

- Mostrar as liñas dun arquivo que contén a palabra PONTEVEDRA e LUGO:

```
$ grep -e "PONTEVEDRA" -e "LUGO" arquivo
```

- Mostrar as liñas dun ficheiro que contén a palabra PONTEVEDRA e LUGO, numerando as liñas de saída:

```
$ grep -n -e "PONTEVEDRA" -e "LUGO" arquivo
```

- Mostrar os ficheiros que conteñen a palabra "LUGO" no directorio actual e todos os seus subdirectorios:

```
$ grep -r "LUGO" *
```

Vexamos algúns exemplos prácticos:

- Obter a dirección MAC da interfaz de rede **eth0** da nosa máquina:

```
$ ifconfig eth0 | grep -oiE '([0-9A-F]{2}:){5}[0-9A-F]{2}'
```

A expresión regular empregada podemos dividila en dúas partes:

- **[0-9A-F]{2}:){5}** Buscamos 5 conxuntos de 2 caracteres seguidos de dous puntos.
- **[0-9A-F]{2}** Seguido por un conxunto de dous caracteres.

Como as direccións MAC se representan en hexadecimal, os caracteres que buscamos son os números do 0 ao 9 e as letras dende a A á F.

- Extraer a lista de direccións de correo electrónico dun arquivo:

```
grep -Eio '[a-z0-9._-]+@[a-z0-9.-]+[a-z]{2,4}' fichero.txt
```

A expresión regular empregada podemos dividila nas seguintes partes:

- **[a-z0-9._-]+** Unha combinación de letras, números, e/ou os símbolos ., _ e - de un ou mais caracteres.
- **@** Seguido dunha arroba.
- **[a-z0-9.-]+** Seguido dunha cadea de letras, números e/ou os símbolos . e -.
- **[a-z]{2,4}** Seguido dunha cadea de entre dous e catro caracteres.

- Obter a dirección IP da interfaz de rede **eth1** da nosa máquina:

```
$ ifconfig eth1 | grep -oiE '([0-9]{1,3}\.){3}[0-9]{1,3}' | grep -v 255
```

A expresión regular empregada podemos dividila nas seguintes partes:

- **([0-9]{1,3}\.){3}** Representa 3 bloques de entre un e tres díxitos separados por puntos. Observar que, como o punto é un metacaracter, hai que empregar o caracter de escape \ para que non sexa interpretado como un metacaracter, sinon como un caracter normal.
- **[0-9]{1,3}** Representa o último bloque da dirección IP, que está formado por un número de entre 1 y 3 díxitos.