

# THAM KHẢO NHANH GREP

Khớp mẫu, regex, tìm kiếm đệ quy, ngữ cảnh, lọc

## Cách Dùng Cơ Bản

### Chạy grep

```
grep "pattern" file.txt # search in file
grep "error" *.log # search multiple files
grep "hello" file1.txt file2.txt # explicit file list
cat file.txt | grep 'pattern' # pipe input
dmesg | grep -i "usb" # filter command output
```

### Cờ Phổ Biến

- i** Khớp không phân biệt hoa thường
- v** Đảo ngược khớp — in dòng không khớp
- c** In số dòng khớp
- n** Hiện thị số dòng
- l** Chỉ liệt kê tên file có khớp
- L** Liệt kê tên file không có khớp
- w** Chỉ khớp toàn bộ từ
- x** Chỉ khớp toàn bộ dòng

## Mẫu Regex

### Biểu Thức Chính Quy Cơ Bản (BRE)

- .** Bất kỳ ký tự đơn nao
- \*** Không hoặc nhiều lần của phần tử trước
- ^** Đầu dòng
- \$** Cuối dòng
- [abc]** Lớp ký tự — bất kỳ trong a, b, c
- ^abc** Lớp phủ định — bất kỳ trừ a, b, c
- [a-z]** Phạm vi — chữ thường
- \<, \>** Ranh giới từ (GNU)
- \( \), \1** Nhóm bắt và tham chiếu ngược

### Ví Dụ BRE

```
grep '^#' file.conf # lines starting with #
grep 'errors$' file.log # lines ending with error
grep '^#' file.txt # blank lines
grep 'col[ou]r' file.txt # match color or colour
```

## Regex Mở Rộng

### Biểu Thức Chính Quy Mở Rộng (ERE)

- +** Một hoặc nhiều lần của phần tử trước
- ?** Không hoặc một lần của phần tử trước
- {n}** Chính xác n lần lặp
- {n,m}** Từ n đến m lần lặp
- (a|b)** Luân phiên — khớp a hoặc b
- ()** Nhóm (không cần dấu gạch chéo ngược)

### Ví Dụ ERE

```
grep -E '[0-9]{3}[0-9]{4}' f # phone number pattern
grep -E '(error|warn|fatal)' f # multiple patterns
grep -E '[A-Z][a-z]+' f # capitalized words
grep -P 'd{1,3}\.d{1,3}' f # Perl regex: IP fragments
```

## Dòng Ngữ Cảnh

### Ví Dụ Ngữ Cảnh

```
grep -B 3 "error" app.log # 3 lines before match
grep -A 5 "FAIL" test.log # 5 lines after match
grep -C 2 "crash" kern.log # 2 lines before and after
grep --group-separator="---" -C 1 "err" f # custom separator
```

### Cờ Ngữ Cảnh

- B N** Hiện thị N dòng trước mỗi khớp
- A N** Hiện thị N dòng sau mỗi khớp
- C N** Hiện thị N dòng trước và sau (ngữ cảnh)
- group-separator=str** Dấu phân cách giữa các nhóm khớp (mặc định: ---)
- color=auto** Tô sáng khớp trong terminal

## Tìm Kiếm Đệ Quy

### Ví Dụ Đệ Quy

```
grep -r "TODO" . # recursive from current dir
grep -rn "FIXME" src/ # recursive with line numbers
grep -r --include="*.py" "import" . # only .py files
grep -r --exclude="*.log" "error" . # skip .log files
grep -r --exclude-dir=node_modules "require" .
```

### Cờ Đệ Quy

- r / --recursive** Tìm kiếm thư mục đệ quy
- R** Như **-r** nhưng theo đối symlink
- include=glob** Chỉ tìm file khớp glob
- exclude=glob** Bỏ qua file khớp glob
- exclude-dir=dir** Bỏ qua thư mục khớp tên
- include-dir=dir** Chỉ tìm trong thư mục khớp tên

## Đếm & Liệt Kê

### Ví Dụ Đếm & Liệt Kê

```
grep -c "error" *.log # count matches per file
grep -l "TODO" src/*.py # list files with TODOs
grep -L "test" src/*.py # files missing "test"
grep -o "http[^"]*" page.html # extract matching parts only
grep -c '' file.txt # count total lines (like wc -l)
```

## Cờ Output

- c** In số dòng khớp mỗi file
- l** Chỉ in tên file có khớp
- L** Chỉ in tên file không có khớp
- o** Chỉ in phần khớp của dòng
- H / -h** Hiện thị / ẩn tiền tố tên file
- Z** Output phân cách null (cho xargs -0)

## Đảo Ngược Khớp

### Đảo Ngược & Loại Trừ

```
grep -v '^#' config.conf # remove comment lines
grep -v '^$' file.txt # remove blank lines
grep -v -e "debug" -e "trace" app.log # exclude two patterns
grep -v "pattern" f | grep "other" # chain: NOT A, then B
```

## Chiến Lược Lọc

### **-v**

Đảo ngược khớp — chọn dòng không khớp

### **-v with -e**

Loại trừ nhiều mẫu

### **pipe chain**

Nối các lệnh grep cho lọc phức tạp

```
grep -v '^$' | grep -v '^#' Xóa dòng trống và comment
```

### **-v with -c**

Đếm dòng không khớp

## Nhiều Mẫu

### Ví Dụ Nhiều Mẫu

```
grep -e "error" -e "warning" app.log
grep -E "error|warning|fatal" app.log
grep -f patterns.txt file.txt # patterns from file
grep -w -e "GET" -e "POST" access.log
```

## Tùy Chọn Mẫu

- e pattnern** Chỉ định mẫu (dùng nhiều lần)
- f file** Đọc mẫu từ file (mỗi dòng một mẫu)
- E 'a|b|c'** Luân phiên ERE cho nhiều mẫu
- F** Chuỗi cố định — không regex, khớp nhanh hơn
- G** Regex cơ bản (chế độ mặc định)
- P** Regex tương thích Perl (PCRE)

## Hiệu Năng

### Mẹo Hiệu Năng

```
-F (fgrep) Chế độ chuỗi cố định — nhanh nhất cho chuỗi literal
```

```
LC_ALL=C grep Bỏ qua locale để tăng tốc 2-10x trên dữ liệu ASCII
```

```
--include/--exclude Giảm số file tìm kiếm trước khi mở
```

```
-m N Dừng sau N khớp mỗi file
```

```
-q Chế độ im lặng — thoát khi tìm thấy khớp đầu tiên (cho script)
```

```
zipgrep (rg) Thay thế drop-in; nhanh hơn trên repo lớn
```

### Ví Dụ Hiệu Năng

```
LC_ALL=C grep -F "exact string" huge.log
grep -r -m 1 "needle" /var/log/ # stop after first hit
grep -rq "pattern" . && echo "found" # boolean test
grep -r --include="*.go" "func main" .
```

## Mẫu Phổ Biến

### One-Liners

```
grep -rn "TODO|FIXME|HACK" src/ # find code markers
grep -oP '(?<=)("[^"]+")' f # extract quoted strings
grep -E '\s*$' f | wc -l # count blank lines
grep -c '' *.py | sort -t: -k2 -rn # sort files by line count
grep -rn --include="*.yaml" "password" . # audit for secrets
```

## Công Thức

```
IP addresses grep -oE '[0-9]{1,3}(\.[0-9]{1,3}){3}'
```

```
Email addresses grep -oE '[a-zA-Z0-9_%+]+@[a-z-]+'
```

```
URLs grep -oE 'https://[^"]+'
```

```
Lines between markers grep -A999 'START' f | grep -B999 'END'
```

```
Unique matches grep -oE 'pattern' f | sort -u
```

```
Count per pattnern grep -c 'pat1' f; grep -c 'pat2' f
```