



链滴

Shell cut 命令：查找符合条件的行

作者：[Leif160519](#)

原文链接：<https://ld246.com/article/1587368692311>

来源网站：[链滴](#)

许可协议：[署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)



grep 命令是在文件中提取符合条件的行，也就是分析一行的信息，如果行中包含需要的信息，就把行提取出来。而如果要进行列提取，就要利用 cut 命令了。

不过要小心，虽然 cut 命令用于提取符合条件的列，但是也要一行一行地进行数据提取。也就是说，要读取文本的第一行数据，在此行中判断是否有符合条件的字段，然后再处理第二行数据。我们也可把 cut 成为字段提取命令。命令格式如下：

cut [选项] 文件名

选项：

- -f 列号：提取第几列；
- -d 分隔符：按照指定分隔符分割列；
- -c 字符范围：不依赖分隔符来区分别，而是通过字符范围（行首为 0）来进行字段提取。"n-"表示第 n 个字符到行尾；"n-m"表示从第 n 个字符到第 m 个字符；"-m"表示从第 1 个字符到第 m 个字；

cut 命令的默认分隔符是制表符，也就是 Tab 键，不想对空格符支持得不怎么好。我们先建立一个测文件，然后看看 cut 命令的作用。

```
cat student.txt
ID Name gender Mark
1 Liming M 86
2 Sc M 90
3 Gao M 83
```

建立学员成绩表，注意这张表中所有的分隔符都是制表符，不能是空格，否则后面的实验会出现问题。

先看看 cut 命令该如何使用：

```
cut -f 2 student.txt
```

#提取第二列的内容

```
Name  
Liming  
Sc  
Gao
```

如果想要提取多列呢？将列号直接用","隔开，命令如下：

```
cut -f 2,3 student.txt  
#提取第二列和第三列的内容  
Name gender  
Liming M  
Sc M  
Gao M
```

cut 命令可以按照字符进行提取。需要注意的是，"8-"代表提取所有行从第 8 个字符到行尾，而"10-2"代表提取所有行的第 10~20 个字符，而"-8"代表提取所有行从行首到第 8 个字符，命令如下：

```
cut -c 8- student.txt  
#提取取每行从第8个字符到行尾，好像很乱啊，那是因为每行的字符个数不相等  
gender Mark  
g M 86  
90  
83
```

当然，cut 命令也可以手工指定分隔符。例如，我想看看当前 Linux 服务器中有哪些用户、这些用户的 UID 是什么，就可以这样操作：

```
cut -d ":" -f 1,3 /etc/passwd  
#以":"作为分隔符，提取/etc/passwd文件的第一列和第三列  
root:0  
bin:1  
daemon:2  
adm:3  
lp:4  
sync:5  
shutdown:6  
halt:7  
mail:8  
operator:11  
games:12  
ftp:14  
nobody:99  
systemd-network:192  
dbus:81  
polkitd:999  
sshd:74  
postfix:89  
chrony:998  
ntp:38  
nginx:997  
rabbitmq:996  
redis:995  
mongod:994  
sasauth:993
```

```
mailnull:47
smmsp:51
zabbix:992
tcpdump:72
```

扩展:可以用`awk -F: '{print $1 ":" $3}' /etc/passwd`同样达到效果

`cut` 命令很方便, 不过最主要的问题是对空格识别得不好, 很多命令的输出格式中都不是制表符, 而空格符, 比如:

```
df -h
文件系统          容量 已用 可用 已用% 挂载点
/dev/mapper/centos_192-root 50G 12G 39G 23% /
devtmpfs          3.8G  0 3.8G  0% /dev
tmpfs             3.9G  0 3.9G  0% /dev/shm
tmpfs             3.9G 269M 3.6G  7% /run
tmpfs             3.9G  0 3.9G  0% /sys/fs/cgroup
/dev/sda1         1014M 173M 842M 18% /boot
/dev/mapper/centos_192-home 42G 33M 42G  1% /home
tmpfs            781M  0 781M  0% /run/user/0
```

如果想用 `cut` 命令截取第一列和第三列, 就会出现这样的情况:

```
df -h | cut -d " " -f 1,3
文件系统
/dev/mapper/centos_192-root
devtmpfs
tmpfs
tmpfs
tmpfs
/dev/sda1
/dev/mapper/centos_192-home
tmpfs
overlay
overlay
overlay
shm
shm
shm
overlay
shm
```

第三列去哪里了? 其实因为 `df` 命令输出的分隔符不是制表符, 而是多个空格符, 所以 `cut` 命令会忠实地将每个空格符当作一个分隔符, 而这样数, 第三列刚好也是空格, 所以输出才会是上面这种情况。

总之, `cut` 命令不能很好地识别空格符。如果想要以空格符作为分隔符, 建议使用 `awk` 命令。

参考:

- [Shell cut命令: 查找符合条件的行](#)
- [Linux cut命令](#)