

cut

连接文件并打印到标准输出设备上

补充说明

cut 命令用来显示行中的指定部分，删除文件中指定字段。cut 经常用来显示文件的内容，类似于 type 命令。

说明：该命令有两项功能，其一是用来显示文件的内容，它依次读取由参数 file 所指明的文件，将它们的内容输出到标准输出上；其二是连接两个或多个文件，如 `cut f1 f2 > f3` 将把文件 f1 和 f2 的内容合并起来，然后通过输出重定向符 “>” 的作用，将它们放入文件 f3 中。

当文件较大时，文本在屏幕上迅速闪过（滚屏），用户往往看不清所显示的内容。因此，一般用 more 等命令分屏显示。为了控制滚屏，可以按 Ctrl+S 键，停止滚屏；按 Ctrl+Q 键可以恢复滚屏。按 Ctrl+C（中断）键可以终止该命令的执行，并且返回 Shell 提示符状态。

语法

```
cut [选项] (参数)
```

选项

- b 仅显示行中指定直接范围的内容；
- c 仅显示行中指定范围的字符；
- d 指定字段的分隔符，默认的分隔符为“TAB”
- f 显示指定字段的内容；
- n 与“-b”选项连用，不分割多字节字符；
- complement 补足被选择的字节、字符或字段；
- out-delimiter= 字段分隔符：指定输出内容是的字段分割符；
- help 显示指令的帮助信息；
- version 显示指令的版本信息。

参数

文件：指定要进行内容过滤的文件。

实例

例如有一个学生报表信息，包含 No Name Mark Percent

```
[root@localhost text]# cat test.txt
No Name Mark Percent
01 tom 69 91
02 jack 71 87
03 alex 68 98
```

使用 -f 选项提取指定字段（这里的 f 参数可以简单记忆为 --fields 的缩写）：

```
[root@localhost text]# cut -f 1 test.txt
No
```

```
01
02
03
```

```
[root@localhost text]# cut -f2,3 test.txt
Name Mark
tom 69
jack 71
alex 68
```

--complement 选项提取指定字段之外的列（打印除了第二列之外的列）：

```
[root@localhost text]# cut -f2 --complement test.txt
No Mark Percent
01 69 91
02 71 87
03 68 98
```

使用 **-d** 选项指定字段分隔符：

```
[root@localhost text]# cat test2.txt
No;Name;Mark;Percent
01;tom;69;91
02;jack;71;87
03;alex;68;98
```

```
[root@localhost text]# cut -f2 -d";" test2.txt
Name
tom
jack
alex
```

指定字段的字符或者字节范围

cut 命令可以将一串字符作为列来显示，字符字段的记法：

- **N-**：从第 N 个字节、字符、字段到结尾；
- **N-M**：从第 N 个字节、字符、字段到第 M 个（包括 M 在内）字节、字符、字段；
- **-M**：从第 1 个字节、字符、字段到第 M 个（包括 M 在内）字节、字符、字段。

上面是记法，结合下面选项将摸个范围的字节、字符指定为字段：

- **-b** 表示字节；
- **-c** 表示字符；
- **-f** 表示定义字段。

示例

```
[root@localhost text]# cat test.txt
abcdefghijklmnopqrstuvwxyz
abcdefghijklmnopqrstuvwxyz
abcdefghijklmnopqrstuvwxyz
```

```
abcdefghijklmnopqrstuvwxyz  
abcdefghijklmnopqrstuvwxyz
```

打印第1个到第3个字符：

```
[root@localhost text]# cut -c1-3 test.txt  
abc  
abc  
abc  
abc  
abc
```

打印前2个字符：

```
[root@localhost text]# cut -c-2 test.txt  
ab  
ab  
ab  
ab  
ab
```

打印从第5个字符开始到结尾：

```
[root@localhost text]# cut -c5- test.txt  
efghijklmnopqrstuvwxyz  
efghijklmnopqrstuvwxyz  
efghijklmnopqrstuvwxyz  
efghijklmnopqrstuvwxyz  
efghijklmnopqrstuvwxyz
```

精彩案例

需求:按小时统计接口(/api/a)的访问次数

文件:内容如下

```
#filename access.irust.top  
irust.top 149.7.38.xx 30.001s - [10/Feb/2021:09:54:42 +0800] "POST /api/a  
HTTP/1.1" 499 0 "-" - "SUP=- SUBP=-" "REQUEST_ID=-" "-" || Ali_Ecom14y110  
irust.top 149.7.38.xx 30.002s - [10/Feb/2021:09:54:45 +0800] "POST /api/a  
HTTP/1.1" 499 0 "-" - "SUP=- SUBP=-" "REQUEST_ID=-" "-" || Ali_Ecom13t186  
irust.top 149.7.38.xx 30.000s - [10/Feb/2021:10:55:15 +0800] "POST /api/a  
HTTP/1.1" 499 0 "-" - "SUP=- SUBP=-" "REQUEST_ID=-" "-" || Ali_Ecom10y138  
irust.top 149.7.38.xx 30.002s - [10/Feb/2021:12:55:12 +0800] "POST /api/a  
HTTP/1.1" 499 0 "-" - "SUP=- SUBP=-" "REQUEST_ID=-" "-" || Ali_Ecom15t140  
....
```

cut 默认以空格为分隔符,所以时间段在第5段

```
cut -d ' ' -f 5
```

```
#output
```

```
[10/Feb/2021:09:54:42  
[10/Feb/2021:09:54:45  
[10/Feb/2021:10:55:15  
[10/Feb/2021:12:55:12
```

再利用 cut 拆分时间段到小时

```
cut -d ':' -f 2
```

```
#output
```

```
09  
09  
10  
12
```

以上合并成一句shell

```
#按小时排序
```

```
cat access.irust.top | cut -d ' ' -f 5|cut -d ':' -f 2|sort |uniq -c
```

```
#按访问量排序
```

```
cat access.irust.top | cut -d ' ' -f 5|cut -d ':' -f 2|sort |uniq -c|sort -  
nr
```

From:

<https://rd.irust.top/> - 学习笔记

Permanent link:

<https://rd.irust.top/doku.php?id=command:cut>

Last update: **2021/10/15 14:58**

