



链滴

linux shell 之 字符处理命令一

作者: [c2y2](#)

原文链接: <https://ld246.com/article/1395749878007>

来源网站: [链滴](#)

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

<p>
1, cut 截取命令

</p>

<p>

cut : 主要用于对行数据的分解

</p>

<p>

常用语法:

</p>

<pre class="prettyprint lang-bsh">用法: cut [选项]... [文件]...

从每个文件中输出指定部分到标准输出。

长选项必须使用的参数对于短选项时也是必需使用的。

-b, --bytes=列表 只选中指定的这些字节
-c, --characters=列表 只选中指定的这些字符#如: cat /etc/p -d, --delimite
=分界符 使用指定分界符代替制表符作为区域分界
-f, --fields=列表 只选中指定的这些域; 并打印所有不包含分界符的
行, 除非-s 选项被指定
-n (忽略)
--complement 补全选中的字节、字符或域
-s, --only-delimited 不打印没有包含分界符的行
--output-delimiter=字符串 使用指定的字符串作为输出分界符, 默认采用输入
的分界符
--help 显示此帮助信息并退出
--version 显示版本信息并退出

<p>仅使用 f -b, -c 或-f 中的一个。每一个列表都是专门为一个类别作出的, 或者您可以用逗号隔

> 开要同时显示的不同类别。您的输入顺序将作为读取顺序, 每个仅能输入一次。

每种参数格式表示范围如下:

N 从第 1 个开始数的第 N 个字节、字符或域

N- 从第 N 个开始到所在行结束的所有字符、字节或域

N-M 从第 N 个开始到第 M 个之间(包括第 M 个)的所有字符、字节或域

-M 从第 1 个开始到第 M 个之间(包括第 M 个)的所有字符、字节或域</p>

<p>当没有文件参数, 或者文件不存在时, 从标准输入读取

</p></pre><p></p>

<p>

例子:

</p>

<p>

1, 截取第一列的内容如

</p>

<p>

</p>

<pre class="prettyprint lang-bsh">root@c2y2:/# cat passwd.txt

root:x:0:0:root:/root:/bin/bash

daemon:x:1:1:daemon:/usr/sbin:/bin/sh

bin:x:2:2:bin:/bin:/bin/sh

sys:x:3:3:sys:/dev:/bin/sh

sync:x:4:65534:sync:/bin:/bin/sync

games:x:5:60:games:/usr/games:/bin/sh

man:x:6:12:man:/var/cache/man:/bin/sh

lp:x:7:7:lp:/var/spool/lpd:/bin/sh

mail:x:8:8:mail:/var/mail:/bin/sh

news:x:9:9:news:/var/spool/news:/bin/sh

```
uucp:x:10:10:uucp:/var/spool/uucp:/bin/sh
list:x:38:38:Mailing List Manager:/var/list:/bin/sh
irc:x:39:39:ircd:/var/run/ircd:/bin/sh
gnats:x:41:41:Gnats Bug-Reporting System (admin):/var/lib/gnats:/bin/sh
nobody:x:65534:65534:nobody:/nonexistent:/bin/sh
root@c2y2:/# cat passwd.txt|cut -d ":" -f 1
root
daemon
bin
sys
sync
games
man
lp
mail
news
uucp
list
irc
gnats
nobody
</pre>
<p>
  <br>
</p>
<p>
  <strong>2,</strong> <b> <span> <span> <strong>sort</strong> </span> </span> </b> <strong>排序命令</strong>
</p>
<p>
  sort:常用于对文本按行进行排序
</p>
<p>
  常用语法:
</p>
<pre class="prettyprint lang-bsh">用法: sort [选项]... [文件]...
或: sort [选项]... --files0-from=F
串联排序所有指定文件并将结果写到标准输出。
<p>长选项必须使用的参数对于短选项时也是必需使用的。 <br>
排序选项: </p>
<p>-b, --ignore-leading-blanks 忽略前导的空白区域<br>
-d, --dictionary-order 只考虑空白区域和字母字符<br>
-f, --ignore-case 忽略字母大小写<br>
-g, --general-numeric-sort 按照常规数值排序<br>
-i, --ignore-nonprinting 只排序可打印字符<br>
-M, --month-sort 比较(未知) &lt; "一月" &lt; ... &lt; "十二月" <br>
在 LC_ALL=C 时为(unknown) &lt; <code>JAN' &amp;lt; ... &amp;lt; </code> DEC' <br>
-h, --human-numeric-sort 使用易读性数字(例如: 2K 1G)<br>
-n, --numeric-sort 根据字符串数值比较<br>
-R, --random-sort 根据随机 hash 排序<br>
--random-source=文件 从指定文件中获得随机字节<br>
-r, --reverse 逆序输出排序结果<br>
--sort=WORD 按照 WORD 指定的格式排序: <br>
一般数字-g, 高可读性-h, 月份-M, 数字-n, <br>
```

随机-R, 版本-V

-V, --version-sort 在文本内进行自然版本排序</p>

<p>其他选项: </p>

```
<pre> <code class="highlight-chroma"> <span class="highlight-line"> <span class="highlight-cl"> --batch-size=NMERGE 一次最多合并NMERGE 个输入; 如果输入更多
</span> </span> <span class="highlight-line"> <span class="highlight-cl"> 则使用
时文件
</span> </span> </code> </pre>
```

<p>-c, --check, --check=diagnose-first 检查输入是否已排序, 若已有序则不进行操作

-C, --check=quiet, --check=silent 类似-c, 但不报告第一个无序行

--compress-program=程序 使用指定程序压缩临时文件; 使用该程序

的-d 参数解压缩文件

--debug 为用于排序的行添加注释, 并将有可能有问题的

用法输出到标准错误输出

--files0-from=文件 从指定文件读取以 NUL 终止的名称, 如果该文件被

指定为 "-" 则从标准输入读文件名

-k, --key=位置 1[位置 2] 在位置 1 开始一个 key, 在位置 2 终止(默认为行尾)

参看 POS 语法。

-m, --merge 合并已排序的文件, 不再进行排序

-o, --output=文件 将结果写入到文件而非标准输出

-s, --stable 禁用 last-resort 比较以稳定比较算法

-S, --buffer-size=大小 指定主内存缓存大小

-t, --field-separator=分隔符 使用指定的分隔符代替非空格到空格的转换

-T, --temporary-directory=目录 使用指定目录而非 \$TMPDIR 或/tmp 作为

临时目录, 可用多个选项指定多个目录

--parallel=N 将同时运行的排序数改变为 N

-u, --unique 配合-c, 严格校验排序; 不配合-c, 则只输出一次排序结果

-z, --zero-terminated 以 0 字节而非新行作为行尾标志

--help 显示此帮助信息并退出

--version 显示版本信息并退出

</p> </pre> <p> </p>

<p>

例子:

</p>

<p>

1, 统计网络访问并按ip排序

</p>

<p>

</p>

```
<pre class="prettyprint lang-bsh"> root@c2y2:/# netstat -atlnup|awk '{print $5}'|sort -n
```

```
...*
```

```
...*
```

```
...*
```

```
...*
```

```
...*
```

```
...*
```

```
0.0.0.0:*
```

```
0.0.0.0:*
```

```
0.0.0.0:*
```

```
0.0.0.0:*
```

```
0.0.0.0:*
```

```
0.0.0.0:*
```

```
Address
```

```
and
```

```
59.188.77.95:80
59.188.77.95:9998
63.245.216.134:443
63.245.217.207:443
74.125.235.198:443
115.239.210.151:80
173.194.72.113:443
173.194.72.113:443
178.63.49.5:443
178.63.49.5:443
220.181.111.96:80</pre>
```

<p>

3,wc 统计命令

</p>

<p>

wc:常用于统计行数, 字节, 或者字符数

</p>

<p>

常用语法:

</p>

```
<pre class="prettyprint lang-bsh">用法: wc [选项]... [文件]...
```

```
-c, --bytes          打印出字节总数
```

```
-m, --chars          打印出字符总数
```

```
-l, --lines          打印出总行数
```

```
--files0-from=文件 从指定文件读取以NUL 终止的名称, 如果该文件被指定为 "-" 则从标准输入读文件名
```

```
-L, --max-line-length 显示最长行的长度
```

```
-w, --words          显示单词计数
```

```
--help              显示此帮助信息并退出
```

```
--version           显示版本信息并退出
```

```
</pre>
```

<p>

例子: 用于统计某个pass.txt的行数

</p>

```
<pre class="prettyprint lang-bsh">root@c2y2:/# cat passwd.txt
```

```
root:x:0:0:root:/root:/bin/bash
```

```
daemon:x:1:1:daemon:/usr/sbin:/bin/sh
```

```
bin:x:2:2:bin:/bin:/bin/sh
```

```
sys:x:3:3:sys:/dev:/bin/sh
```

```
sync:x:4:65534:sync:/bin:/bin/sync
```

```
games:x:5:60:games:/usr/games:/bin/sh
```

```
man:x:6:12:man:/var/cache/man:/bin/sh
```

```
lp:x:7:7:lp:/var/spool/lpd:/bin/sh
```

```
mail:x:8:8:mail:/var/mail:/bin/sh
```

```
news:x:9:9:news:/var/spool/news:/bin/sh
```

```
uucp:x:10:10:uucp:/var/spool/uucp:/bin/sh
```

```
list:x:38:38:Mailing List Manager:/var/list:/bin/sh
```

```
irc:x:39:39:ircd:/var/run/ircd:/bin/sh
```

```
gnats:x:41:41:Gnats Bug-Reporting System (admin):/var/lib/gnats:/bin/sh
```

```
nobody:x:65534:65534:nobody:/nonexistent:/bin/sh
```

```
root@c2y2:/# cat passwd.txt |wc -l
```

```
15</pre>
```

<p>

 4, uniq 合并匹配命令

</p>

<p>

uniq: 这个指令用来将『重复的行删除掉只显示一个』（相邻的行才有效哦）

</p>

<p>

常用语法:

</p>

<pre class="prettyprint lang-bsh">用法: uniq [选项]... [文件]

从输入文件或者标准输入中筛选相邻的匹配行并写入到输出文件或标准输出。

不附加任何选项时匹配行将在首次出现处被合并。

长选项必须使用的参数对于短选项时也是必需使用的。

- c, --count 在每行前加上表示相应行目出现次数的前缀编号
- d, --repeated 只输出重复的行
- D, --all-repeated[=delimit-method 显示所有重复的行
delimit-method={none(default),prepend,separate}
以空行为界限
- f, --skip-fields=N 比较时跳过前N列
- i, --ignore-case 在比较的时候不区分大小写
- s, --skip-chars=N 比较时跳过前N个字符
- u, --unique 只显示唯一的行
- z, --zero-terminated 使用'\0'作为行结束符,而不是新换行
- w, --check-chars=N 对每行第N个字符以后的内容不作对照
- help 显示此帮助信息并退出
- version 显示版本信息并退出

</pre>

例子: 查看网络链接并合并同ip的

<pre class="prettyprint lang-bsh">root@c2y2:/# netstat -atlnup

Active Internet connections (servers and established)

| Proto | Recv-Q | Send-Q | Local Address | Foreign Address | State | PID/Program name |
|-------|--------|--------|-------------------------|--------------------|-------------|--------------------|
| tcp | 0 | 0 | 192.168.28.126:34782 | 59.200.33.95:80 | ESTABLISHED | 3304/google-chrome |
| tcp | 0 | 0 | 192.168.28.126:43808 | 74.125.31.100:443 | ESTABLISHED | 3304/google-chrome |
| tcp | 0 | 0 | 192.168.28.126:35103 | 63.245.217.207:443 | ESTABLISHED | 11524/iceweasel |
| tcp | 0 | 0 | 192.168.28.126:52287 | 59.200.33.95:9998 | ESTABLISHED | 12273/fuckyou |
| tcp | 0 | 0 | 192.168.28.126:49368 | 74.125.235.198:443 | ESTABLISHED | 3304/google-chrome |
| tcp | 0 | 0 | 192.168.28.126:52504 | 178.63.49.5:443 | ESTABLISHED | 11524/iceweasel |
| tcp | 0 | 0 | 192.168.28.126:52499 | 178.63.49.5:443 | TIME_WAIT | - |
| tcp | 0 | 0 | 192.168.28.126:51084 | 63.245.216.134:443 | ESTABLISHED | 11524/iceweasel |
| udp | 0 | 0 | 0.0.0.0:32824 | 0.0.0.0:* | | 7389/dhclient |
| udp | 0 | 0 | 0.0.0.0:68 | 0.0.0.0:* | | 7389/dhclient |
| udp | 0 | 0 | 192.168.1.105:123 | 0.0.0.0:* | | 2499/ntpd |
| udp | 0 | 0 | 192.168.28.126:123 | 0.0.0.0:* | | 2499/ntpd |
| udp | 0 | 0 | 127.0.0.1:123 | 0.0.0.0:* | | 2499/ntpd |
| udp | 0 | 0 | 0.0.0.0:123 | 0.0.0.0:* | | 2499/ntpd |
| udp6 | 0 | 0 | :::64957 | :::* | | 7389/dhclient |
| udp6 | 0 | 0 | fe80::222:5fff:febf:123 | :::* | | 2499/ntpd |
| udp6 | 0 | 0 | fe80::182b:3cff:fe4:123 | :::* | | 2499/ntpd |
| udp6 | 0 | 0 | :::1:123 | :::* | | 2499/ntpd |
| udp6 | 0 | 0 | :::123 | :::* | | 2499/ntpd |

root@c2y2:/# netstat -atlnup|awk '{print \$5}'|sort|uniq -c

5 :::*

```
6 0.0.0.0:*
2 178.63.49.5:443
1 59.200.33.95:80
1 59.200.33.95:9998
1 63.245.216.134:443
1 63.245.217.207:443
1 74.125.235.198:443
1 74.125.31.100:443
1 Address
1 and
```

</pre>

<p>

</p>

<p>

后记：例子没几个， :-)，使用起来很简单，使用时直接看其使用手册就ok拉， man command ^_。
。男人嘛，多man才行哦！

</p>

<p>

</p>