常见正则表达式实例学习，结合find、grep、后续的awk和sed

SED、AWK学习

SED通过管道（|）和AWK配合使用。

Linux之父：找份Linux工作是明智选择

  
1991年， 也就是25年前托瓦兹以个人兴趣爱好启动了Linux的开发， 他想要打造一款满足他需求的操作系统。时至今日， Linux俨然成为了数以千计GNU/Linux操作系统和项目的幕后功臣。  
Linux已经涉及日常生活的方方面面， 驱动互联网的大型服务器、 汽车、 移动手机和部分嵌入式设备等都能够看到Linux的身影。 但是更为重要的是Linux并非都是枯燥的技术， 在托瓦兹看来背后的Linux社区涌现了很多富有创意的开源项目，为你的未来职业生涯创造了无限可能。 因此托瓦兹推荐你考虑选择在这片领域中发展， 这或许将会成为你这辈子做的最明智决定。 这个世界需要更多的Linux和开源专家。

-----------------------------------

SED删除行方面操作，AWK擅长列方面操作，操作都需要正则表达式来定位。

Linux系统一切皆文件，对Linux的操作就是对文本的操作。针对文本操作，有“Linux三剑客”——grep、sed、awk。其中grep命令擅长查找功能；sed擅长取行和替换；awk擅长取列。这三款工具均可结合正则表达式，达到精确操纵文本，进而操纵Linux系统。

awk、sed、grep可以结合管道、重定向，也可以写在脚本中。

先生成示范文本

rpm -qai >> software-info

rpm -qai | head -15 >sedtest

（1）正则表达式

正则表达式是组成“样式”的基本语法，而“样式”是运用 sed 和 awk 必备的能力。 sed和 awk 相同的运行方式是：只要符合“样式”的数据行，就对它执行指定的“操作”。

结合grep语法

grep -E -o “match\_pattern" file\_name1 file\_name2 file\_name3

-o 表示只输出文件中匹配到的部分

-E 表示使用扩展正则表达式

rpm -qai | head -15 > re-test

（2）SED简明教程

sed语法如下：

sed '样式命令' 文件

sed 的工作方式是读取文件内容，经流编辑之后，把结果显示到标准输出。

sed默认是将结果输出到屏幕上，因此其不会改变文件内容，想要改变文件内容，需要加一个 -i 选项；或者重定向到临时文件，然后确认没问题后，用mv 将临时文件覆盖源文件。

#删除1行到5行数据，并修改到源文件

sed -i '1,5d' sedtest

#删除第3行数据，并修改到源文件

sed -i '3d' sedtest

也可也这么操作：

sed '3d' >tmp

mv tmp sedtest

#将S打头的行，打印出来,-n选项意味着只输出符合条件的

sed -n '/^S/p' sedtest

#删除空白行

sed '/^$/d' sedtest

#将Bulid替换成Make，并打印输出到屏幕

sed -n 's/Build/Make/' sedtest

#将包含2016那一行中的Build替换为Make

sed -n '/2016/s/Build/Make/p' sedtest

#将第一行到第七行，所有的6替换成x，而不是第一个6

[root@localhost jkjk]# sed -n '1,7s/6/x/pg' sedtest

（3）AWK简明教程

awk '样式'文件：把符合样式的数据行显示出来。

awk '{操作}' 文件：对每一行都执行{}中的操作。

awk '样式{操作}' 文件：对符合样式的数据行，执行{}中的操作。

#显示URL所在行

awk '/URL/' software\_info

#打印出以：为分割，第二行

awk -F : '{print $2 }' software\_info

cat software\_info | head -20 | awk -F: '{print $2}'

#查找含有La关键字的数据行，中第三，四列，默认列分隔符是空格

awk '/La/{ print $3,$4}' 文件

#将ftp打头的账户1~3栏信息打印出来，以+++作为输出分隔符，以：作为输入分割符

awk -F: 'BEGIN{OFS="++++"} /^ftp/{print $1,$2,$3}' /etc/passwd

##获取网卡IP(方法有很多)

1.ifconfig | grep 'broadcast' | awk '{print $2}'

2.ifconfig | grep inet | awk '{print $2}' | head -1

3.ifconfig | awk '/inet/{print $2}' | head -1

3.hostname -I

（4）grep用法

（5）其他一些命令

sort，uniq

SED AWK案例展示

[root@localhost learn-script]# rpm -qai | head > ak

[root@localhost learn-script]# cat ak

Name : openssh-server

Version : 6.6.1p1

Release : 31.el7

Architecture: x86\_64

Install Date: Thu 05 Oct 2017 02:01:54 PM PDT

Group : System Environment/Daemons

Size : 943606

License : BSD

Signature : RSA/SHA256, Sun 20 Nov 2016 11:48:43 AM PST, Key ID 24c6a8a7f4a80eb5

Source RPM : openssh-6.6.1p1-31.el7.src.rpm

要查找size大小：

----------------------------

有如下方法：

awk -F: '/Size/{print $2}' ak

awk -F: '/Size/' ak | awk -F: '{print $2}'

grep "Size" ak | awk -F: '{print $2}'

[root@localhost learn-script]# cat awktest

姓名,学号,地址,爱好

zhangpp,20150901,福建省福州市仓山区,跑步

zhangp,20150902,福建省泉州市南安市,登山

zhang123,20150903,福建省厦门市湖里区,羽毛球

lilin,20150902,福建省泉州市南安市,登山

chenyi,20150903,福建省厦门市湖里区,骑行

jiajia,20150902,福建省泉州市南安市,篮球

nihao,20150903,福建省厦门市湖里区,羽毛球

###将输出分隔符由逗号变为空格

[root@localhost learn-script]# awk -F, 'BEGIN{OFS=" "}{print $1,$2,$3,$4,$5}' awktest

姓名 学号 地址 爱好

zhangpp 20150901 福建省福州市仓山区 跑步

zhangp 20150902 福建省泉州市南安市 登山

zhang123 20150903 福建省厦门市湖里区 羽毛球

lilin 20150902 福建省泉州市南安市 登山

chenyi 20150903 福建省厦门市湖里区 骑行

jiajia 20150902 福建省泉州市南安市 篮球

nihao 20150903 福建省厦门市湖里区 羽毛球

##此时并没有改变awktest内容，需要将输出屏幕内容打到临时文件，再将临时文件替换awktest，具体如下：

[root@localhost learn-script]# awk -F, 'BEGIN{OFS=" "}{print $1,$2,$3,$4,$5}' awktest > tmp

[root@localhost learn-script]# mv tmp awktest

mv: overwrite ‘awktest’? y

[root@localhost learn-script]# cat awktest

姓名 学号 地址 爱好

zhangpp 20150901 福建省福州市仓山区 跑步

zhangp 20150902 福建省泉州市南安市 登山

zhang123 20150903 福建省厦门市湖里区 羽毛球

lilin 20150902 福建省泉州市南安市 登山

chenyi 20150903 福建省厦门市湖里区 骑行

jiajia 20150902 福建省泉州市南安市 篮球

nihao 20150903 福建省厦门市湖里区 羽毛球

#将学号姓名对调显示，此时不需要指定输入分隔符，因为awk默认分隔符就是空格

[root@localhost learn-script]# awk '{print $2,$1,$3,$4}' awktest

学号 姓名 地址 爱好

20150901 zhangpp 福建省福州市仓山区 跑步

20150902 zhangp 福建省泉州市南安市 登山

20150903 zhang123 福建省厦门市湖里区 羽毛球

20150902 lilin 福建省泉州市南安市 登山

20150903 chenyi 福建省厦门市湖里区 骑行

20150902 jiajia 福建省泉州市南安市 篮球

20150903 nihao 福建省厦门市湖里区 羽毛球

#将爱好信息提取出来，sort用于排序，将相同爱好的放在一起，用uniq删除相同的爱好

[root@localhost learn-script]# awk '{print $4}' awktest | sort | uniq

爱好

登山

篮球

羽毛球

跑步

骑行

#统计各个爱好的人数

[root@localhost learn-script]# awk '{print $4}' awktest | sort | uniq -c

1 爱好

2 登山

1 篮球

2 羽毛球

1 跑步

1 骑行

#统计各个爱好的人数，并将其输出重定向到tmp文件中

[root@localhost learn-script]# awk '{print $4}' awktest | sort | uniq -c > tmp

#输出爱好登山的人数

[root@localhost learn-script]# awk '/登山/{print $1}' tmp

2

#或者

[root@localhost learn-script]# awk '{print $4}' awktest | sort | uniq -c | awk '/登山/{print $1}'

2

---------------------------------------------------------

[root@localhost learn-script]# cat sedtest

姓名,学号,地址,爱好

zhangpp,20150901,福建省福州市仓山区,跑步

lilin,20150902,福建省泉州市南安市,登山

chenyi,20150903,福建省厦门市湖里区,骑行

jiajia,20150902,福建省泉州市南安市,篮球

nihao,20150903,福建省厦门市湖里区,羽毛球羽毛球

chener,20150903,福建省厦门市湖里区,骑行

jiaren,20150902,福建省泉州市南安市,足球

nini,20150903,福建省厦门市湖里区,羽毛球

#删除第2，3行

[root@localhost learn-script]# sed '2,3d' sedtest

姓名,学号,地址,爱好

chenyi,20150903,福建省厦门市湖里区,骑行

jiajia,20150902,福建省泉州市南安市,篮球

nihao,20150903,福建省厦门市湖里区,羽毛球羽毛球

chener,20150903,福建省厦门市湖里区,骑行

jiaren,20150902,福建省泉州市南安市,足球

nini,20150903,福建省厦门市湖里区,羽毛球

#删除第1行

[root@localhost learn-script]# sed '1d' sedtest

zhangpp,20150901,福建省福州市仓山区,跑步

lilin,20150902,福建省泉州市南安市,登山

chenyi,20150903,福建省厦门市湖里区,骑行

jiajia,20150902,福建省泉州市南安市,篮球

nihao,20150903,福建省厦门市湖里区,羽毛球羽毛球

chener,20150903,福建省厦门市湖里区,骑行

jiaren,20150902,福建省泉州市南安市,足球

nini,20150903,福建省厦门市湖里区,羽毛球

#删除2，3两行，并更改到源文件，有两个办法

[root@localhost learn-script]# sed '2,3d' sedtest > tmp

[root@localhost learn-script]# mv tmp sedtest

mv: overwrite ‘sedtest’? y

#或者

sed -i '2,3d' sedtest

#将sedtest中以chen打头的信息显示出来

[root@localhost learn-script]# sed -n '/^chen/p' sedtest

chenyi,20150903,福建省厦门市湖里区,骑行

chener,20150903,福建省厦门市湖里区,骑行

#如果不加-n，则显示如下

[root@localhost learn-script]# sed '/^chen/p' sedtest

姓名,学号,地址,爱好

chenyi,20150903,福建省厦门市湖里区,骑行

chenyi,20150903,福建省厦门市湖里区,骑行

jiajia,20150902,福建省泉州市南安市,篮球

nihao,20150903,福建省厦门市湖里区,羽毛球羽毛球

chener,20150903,福建省厦门市湖里区,骑行

chener,20150903,福建省厦门市湖里区,骑行

jiaren,20150902,福建省泉州市南安市,足球

nini,20150903,福建省厦门市湖里区,羽毛球

#提取出爱好信息，这个案例中通过管道将sed和awk结合在一起

[root@localhost learn-script]# sed -n '/^chen/p' sedtest | awk -F, '{print $4}'

骑行

骑行

#将nihao 的爱好羽毛球改为睡觉

[root@localhost learn-script]# sed -n '/nihao/s/羽毛球/睡觉/p' sedtest

nihao,20150903,福建省厦门市湖里区,睡觉羽毛球

[root@localhost learn-script]# sed -n '/nihao/s/羽毛球/睡觉/pg' sedtest

nihao,20150903,福建省厦门市湖里区,睡觉睡觉

[root@localhost learn-script]# date

Tue Mar 27 14:51:22 PDT 2018

[root@localhost learn-script]# date | awk '{print $6$2$1}'

2018MarTue

[root@localhost learn-script]# date | awk '{print $6,$2,$1}'

2018 Mar Tue

[root@localhost learn-script]#