**linux文本处理三剑客**

grep、sed和awk都是文本处理工具，虽然都是文本处理工具单却都有各自的优缺点，一种文本处理命令是不能被另一个完全替换的，否则也不会出现三个文本处理命令了。只不过，相比较而言，sed和awk功能更强大而已，且已独立成一种语言来介绍。

grep：文本过滤器，如果仅仅是过滤文本，可使用grep，其效率要比其他的高很多；

sed：Stream EDitor，流编辑器，默认只处理模式空间，**不处理原数据**，如果你处理的数据是针对行进行处理的，可以使用sed；

awk：报告生成器，格式化以后显示。如果对处理的数据需要生成报告之类的信息，或者你处理的数据是按列进行处理的，最好使用awk。

grep

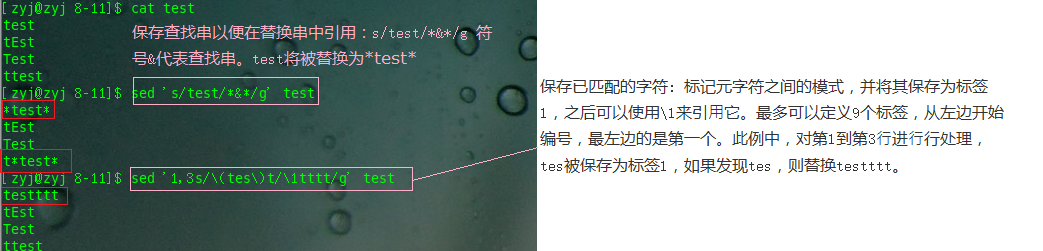
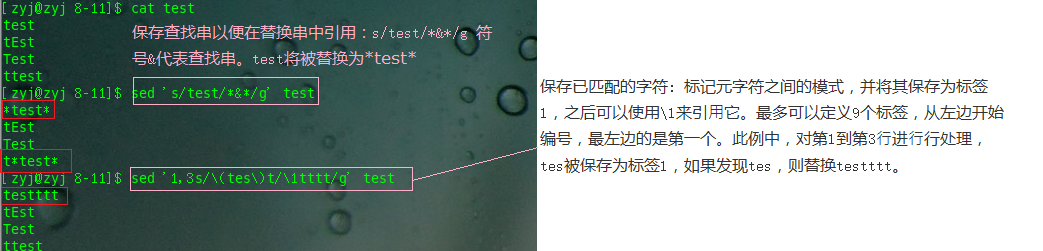
grep(关键字: 截取) 文本搜集工具, 结合正则表达式非常强大  
主要参数 []  
**-c** : 只输出匹配的行  
**-i**: 不区分大小写  
**-h** : 查询多文件时不显示文件名  
**-l** : 查询多文件时, 只输出包含匹配字符的文件名  
**-n** : 显示匹配的行号及行  
**-v** : 显示不包含匹配文本的所有行(我经常用除去grep本身)

基本工作方式: grep 要匹配的内容 文件名, 例如:  
       grep 'test' d\* 显示所有以d开头的文件中包含test的行  
       grep 'test' aa bb cc 显示在 aa bb cc 文件中包含test的行  
       grep '[a-z]\{5}\' aa 显示所有包含字符串至少有5个连续小写字母的串

sed

sed(关键字: 编辑) 以行为单位的文本编辑工具 sed可以直接修改档案, 不过一般不推荐这么做, 可以分析 standard input  
基本工作方式: sed [-nef] '[动作]' [输入文本]  
          a\ ： 在当前行后添加一行或多行。多行时除最后一行外，每行末尾需用“\”续行  
　　    c\ ：用此符号后的新文本替换当前行中的文本。多行时除最后一行外，每⾏末尾需用”\"续行  
　  　  i\ ：在当前行之前插入文本。多行时除最后一行外，每行末尾需用”\"续行删除行  
　　    h ： 把模式空间里的内容复制到暂存缓冲区  
　　    H ： 把模式空间里的内容追加到暂存缓冲区  
　　   g ： 把暂存缓冲区里的内容复制到模式空间，覆盖原有的内容  
　　   G： 把暂存缓冲区的内容追加到模式空间⾥，追加在原有内容的后面  
　   　l ： 列出非打印字符  
　　   p ： 打印行  
　　   q ： 结束或退出sed  
　　   r ： 从文件中读取输入行  
　   　! ： 对所选行以外的所有行应用命令  
　  　 s ： 用一个字符串替换另一个  
　 　  g ： 在行内进行全局替换  
　     w ： 将所选的行写入文件  
　　   x ： 交换暂存缓冲区与模式空间的内容  
　 　  y ： 将字符替换为另一字符（不能对正则表达式使用y命令）  
选项  
　　-e ： 进行多项编辑，即对输入行应用多条sed命令时使用  
　　-n ： 取消默认的输出  
　　-f ：指定sed脚本的文件名

示例：



awk

sed以行为单位处理文件,awk比sed强的地方在于不仅能以行为单位还能以列为单位处理文件。 awk缺省的行分隔符是换行,缺省的列分隔符是连续的空格和Tab,

但是行分隔符和列分隔符都可以自定义,比如/etc/passwd文件的每一行有干个字段,字段之间以:分隔,就可以重新定义awk的列分隔符为:并以列为单位处理这个文件。

awk实际上是一门很复杂的脚本语言,还有像C语言一样的分支和循环结构,但是基本语法和sed类似,awk命令行的基本形式为:

　　awk option 'script' file1 file2 ...

　　awk option -f scriptfile file1 file2 ...  
　　和sed一样,awk处理的文件既可以由标准输入重定向得到,也可以当命令行参数传入,编辑命令可以直接当命令行参数传入,也可以用-f参数指定一个脚本文件,

编辑命令的格式为:

　　/pattern/{actions}

　　和sed类似,pattern是正则表达式,actions是一系列操作。 awk程序一行一行读出待处理文件,如果某一行与pattern匹配,或者满足condition条件,

则执行相应的actions,如果一条awk命令只有actions部分,则actions作用于待处理文件的每一行。

示例：

