霍格沃兹测试学院-测试开发工程师的黄埔军校

三剑客实战Nginx日志分析

MrDong

目录

- 三剑客
- 找出log中的404 500的报错有多少条
- 找出访问量最高的前三名ip
- 替换topics及id横排

正则表达式

普通正则

常用的元字符	
代码	说明
.	匹配除换行符以外的任意字符
\s	匹配任意的空白符
\d	匹配数字
\b	匹配单词的开始或结束
^	匹配字符串的开始
$	匹配字符串的结束
*	表示匹配前边一个字符出现0次或者多次

扩展正则

常用的限定符	
代码/语法	说明
+	重复一次或更多次
?	重复零次或一次
{n}	重复n次
{n,}	重复n次或更多次
{n,m}	重复n到m次
|	表示或

不同版本对比

- POSIX BRE
- POSIX ERE
- GNU BRE
- GNU ERE
- Perl

三剑客

grep

- 查找文件内容包含root的行数
  - grep –n root test.txt
- 查找文件内容不包含root的行
  - grep –nv root test.txt

grep

- 查找以s开头的行
  - grep ^s test.txt
- 查找以n结尾的行
  - grep n$ test.txt

sed

- 在第四行后添加新字符串
  - sed -e '4 a newline testfile' test.txt
- 在第二行后加上 newLine
  - sed '2a drink tea' test.txt
- 在第二行前加上 newline
  - sed '2i drink tea' test.txt

sed

- 全局替换
  - sed -e 's/root/hello/g' test.txt
- 直接修改文件内容
  - sed -i 's/root/hello/g' test.txt

awk

- 搜索/etc/passwd有root关键字的所有行，并显示对应的shell
  - awk -F: '/root/ {print $7}' /etc/passwd
- 打印/etc/passwd/的第二行信息
  - awk -F: 'NR==2{print $0}' /etc/passwd

awk

- 使用begin加入标题
  - awk ‘BEGIN {print “BEGIN” ,“BEGIN” } { print $1,$2 }’ /etc/passwd
- 自定义分割符
  - echo "111 222|333 444|555 666"|awk 'BEGIN{RS="|"}{print $0}'

nginx实战

找出log中的404 500的报错有多少条

- grep -E '\s500\s|\s404\s' nginx.log | wc –l
- awk '$9~/404|500/' nginx.log |wc -l

访问量最高的ip

- awk '{print $1}' nginx.log | sort | uniq -c | sort -nr | head -3
- cat /tmp/nginx.log | grep -o '^[0-9]*.[0-9]*.[0-9]*.[0-9]*' | sort| uniq -c | sort -rn| head -3

将topics后面的数字替换成numer

- grep '/topics/' nginx.log | sed 's@/topics/[0-9]*@/topics/number@'

将ip地址横向打印

- awk '{print $1}' nginx.log | sed ':1;N;s/\n/|/;t1'