**awk基本用法**

awk命令解析

awk工具概述

awk编程语言/数据处理引擎

创造者：Aho、Weinberger、kernighan

基于模式匹配检查输入文本，逐行处理并输出

通常用在shell脚本中，获取指定的数据

单独用时，可对文本数据做统计

命令格式解析

主要用法

格式1：前置命令 | awk [选项] ‘[条件]{指令}’

格式2：awk [选项] ‘[条件]{指令}’ 文件...

常用选项

-F：指定分隔符，可省略（默认空格或TAB位）

常用指令是print；若有多条编辑指令，可用分号以隔开

支持仅打印某一列，如第2列、第5列





指定分隔符

指定多个分隔符



打常量

过滤网卡流量

检测内存剩余量

打印根分区大小

awk内置变量

有特殊含义，可直接使用

|  |  |
| --- | --- |
| 变量 | 用途 |
| FS | 保存或设置字段分割符，例如FS=”:” |
| $n | 指定分隔的第n个字段，如$1、$3分别表示第1、第3列 |
| $0 | 当前读入的整行文本内容 |
| NF | 记录当前处理行的字段个数（列数） |
| NR | 记录当前已读入行的数量（行数） |

awk ‘BEGIN{指令} 条件{指令} END{指令}’ 文件

awk过滤的时机

在所有行前处理，BEGIN{}

读入第一行文本之前执行，执行一次

一般用来初始化操作

逐行处理，{}

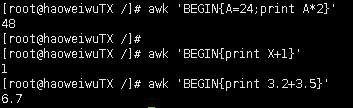
逐行读入文本执行相应的处理，执行n次

是最常见的编辑指令块

在所有行后处理，END{}

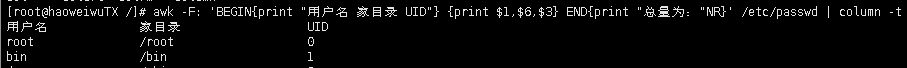
处理完最后一行文本后执行，执行一次

一般用来输出处理结果





变量可以不赋值



输出用户名 家目录 UID 总量，column -t可以排版

awk处理条件

处理条件概述

所有的行全部处理并输出吗？

怎么限制处理条件？

根据多个条件来处理指定的行？

格式回顾：

awk [选项] ‘[条件]{编辑指令}’ 文件....

条件的表现形式：

正则表达式

数值/字符串比较

逻辑比较

运算符

条件设置示例

正则表达式

/正则表达式/

~匹配、!~不匹配











数值和字符串比较

= =等于、!=不等于

>大于、>=大于等于

<小于、<=小于等于

打印第三行

打印uid大于1000的



多个条件组合

逻辑比较测试

&&逻辑与：期望多个条件都成立

||逻辑或：只要有一个条件成立即满足需求





变量的运算

运算符

+-\*/%

++ -- += -= \*= /=

计算200以内能同时被3和13整除的整数个数

**awk高级应用**

awk流程控制

分支结构

单分支：if(条件){编辑指令}

双分支：if(条件){编辑指令1}else{编辑指令2}

多分支：if(条件){编辑指令1}else if(条件){编辑指令2}....

else{编辑指令N}





awk数组

数组的定义及使用

定义数组

格式：数组名[下标]=元素值

调用数组

格式：数组名[下标]

遍历数组

用法：for(变量名 In 数组名){print 数组名[变量]}

