

lxw0109

<https://github.com/lxw0109>

博客园

首页

新随笔

联系

订阅

管理

Efficiency in Shell

最近在写一个shell脚本, 由于该脚本对效率有较高的要求, 所以查阅了一些文章. 感觉这篇文章写得确实不错, 文章中的例子, 确实很棒! 所以我把他们提取出来:

@1:实例:

需求: 计算1到100000累加结果

方法1: 采用bash shell的数值计算

```
# time for((i=0;i<=100000;i++)); do ((sum+=i)); done ; echo $sum
```

```
real    0m1.134s
```

```
user    0m1.080s
```

```
sys     0m0.048s
```

```
5000050000
```

方法2: 采用awk的数值计算

```
# time awk 'BEGIN{while(i++<100000)sum+=i; printf "%d", sum;}'
```

```
5000050000
```

```
real    0m0.029s
```

```
user    0m0.020s
```

```
sys     0m0.000s
```

实例结果分析: 同等的累加运算, 采用shell的数值计算耗时1.134秒, 采用awk工具只需要0.029秒。实验结果表明, 适当的采用工具软件,

能够明显提高shell脚本的执行效率。

@2:实例: 正则效率。

需求: 某文本以空格和/进行分隔, 获取第五列的内容。

方法1: 采用正则[/]分别匹配二种分隔符, 取得第五列

```
# time awk -F'[ /]'{print $5}' a.txt >/dev/null
```

```
real    0m17.717s
```

```
user    0m14.749s
```

```
sys     0m2.844s
```

方法2: 实际数据分析中发现, 可以采用单字符解析的方式, 首先根据“空格”分隔符取得第4列, 再利用“/”分隔符取得第二列。

```
# time awk '{print $4}' a.txt |awk -F/ '{print $2}' >/dev/null
```

```
real    0m0.565s
```

```
user    0m0.224s
```

```
sys     0m0.688s
```

实例结果分析: awk采用-F指定分隔符, 在分隔符情况下, 会启用正则去解析记录, 增加了函数调用, 和字符串匹配的消耗, 效率上理所

当然的比不上默认的空格或单字符分隔符采用的简单的字符比较方式, 尽管如此, 方法2依赖于对需求的进一步分析, 适用的需求范围没有

使用正则那么简便灵活。

@3:实例: sed应用, 读取指定行

方法1: sed -n '45,50p' filename

方法2: sed -n '51q;45,50p' filename

实例结果分析: 方法2在方法1基础上增加了一个判断, 当文件读取到第51行时即时退出。避免文件后续部分的遍历, 在大数据量处理上能

够很大的提高执行效率。

@4:实例: sed应用, 文本替换

方法1: sed 's/foo/bar/g' filename

方法2: sed '/foo/ s/foo/bar/g' filename

实例结果分析: sed支持采用正则进行匹配和替换, 考虑字符串替换的需求中, 不防加上地址以提高速度。实例中通过增加一个判断逻辑,

采用“事先匹配”代替“直接替换”, 由于sed会保留前一次的正则匹配环境, 不会产生冗余的正则匹配, 因此方法2具有比方法1更高

公告

昵称: XiaoweiLiu

园龄: 11年9个月

粉丝: 19

关注: 20

-取消关注

2023年12月						
日	一	二	三	四	五	六
26	27	28	29	30	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1	2	3	4	5	6

最新随笔

- 这样才能使本地Mysql服务允许被外部主机连接 (两步)
- [Python] logging.logger
- Python Selenium
- MySQL中char、varchar和text的区别
- Way to MongoDB
- Python误区之strip,lstrip,rstrip
- Android Studio Tips
- Way to tmux
- Singleton: this & instance
- Java Basis

我的标签

Python(37)
linux(22)
Algorithm(19)
JAVA(17)
shell(16)
NetWork(13)
Study(9)
C++(9)
C#(8)
VIM(4)
更多

积分与排名

积分 - 122622

排名 - 11244

随笔档案 (150)

2017年2月(2)
2017年1月(2)
2016年12月(2)
2016年11月(1)
2016年9月(1)
2016年6月(1)
2016年4月(1)
2016年2月(2)
2015年12月(1)
2015年9月(1)
2015年8月(2)
2015年7月(7)
2015年5月(3)

的效率

@5:awk的特色数据结构：哈希数组，也称关联数组。哈希数组中的每个单元包括二个元素：单元键（key)及单元值（values）。**哈希结构**

的效益体现在以较小的内存空间实现大的数据空间上的数据存放，节省了数据占存空间；同时哈希数组中的数据访问是随机访问，不需要


遍历而直接通过哈希函数直接访问数组单元取值，节省了数据查询时间。

Reference:

shell脚本效率: <http://blog.csdn.net/yanook/article/details/8395482>

标签: [linux](#) , [shell](#)






XiaoweiLiu

粉丝 - 19 关注 - 20





我在关注他 [取消关注](#)

« 上一篇: [Shell Script Practice 2 Summary](#)
» 下一篇: [Simple Tips for Collection in Python](#)


posted @ 2014-08-31 21:49 XiaoweiLiu 阅读(267) 评论(0) 编辑 收藏 举报

 发表评论 升级成为园子VIP会员

编辑 预览

B    

支持 Markdown

 自动补全

[提交评论](#) [退出](#) [订阅评论](#) [我的博客](#)

[Ctrl+Enter快捷键提交]

- 【推荐】编程路上的催化剂：大道至简，给所有人看的编程课
- 【推荐】阿里云热销爆款云服务器，新老同享一口价99元/年
- 【推荐】天翼云S3云主机年度精选低至9.5元/月，立即抢购

编辑推荐:

- 记一次 .NET 某新能源材料检测系统 崩溃分析
- MySQL 分组排序后，如何取前N条或倒数N条
- 「布局技巧」Flex 布局下居中溢出滚动截断问题
- 记一次奇怪的文件句柄泄露问题
- .NET8 极致性能优化 AOT

阅读排行:

- 我的 2023 年，35岁、父亲肺癌，失业，失恋、上岸
- .NET Conf 2023 Chengdu - 成都站圆满结束！
- 基于.NET Core + Quartz.NET+ Vue + IView开箱即用的定时任务UI
- 推荐一款好用的代码可视化工具
- 2023年奔走的总结---吉特日化MES 制药项目 篇二

2015年3月(3)

2015年2月(2)

[更多](#)

阅读排行榜

1. VS中代码对齐等快捷键(38795)
2. 如何解决Protel 99SE添加库文件时出现"File is not recognized"(35276)
3. Java FAQ -- "Exception in thread 'main' java.lang.UnsupportedClassVersionError:"(28961)
4. Python 函数名作为字典值(10904)
5. 父目录的权限对子目录有没有影响? [Linux](10554)

评论排行榜

1. War-ftpd USER longString漏洞攻击之Java实现常见问题(8)
2. C# 静态构造函数和实例构造函数(4)
3. Amazon2014在线笔试 第三题(3)
4. 鼠标滚轮实现窗体图形缩放MouseWheel(3)
5. 程序设计语言的顺序点问题（参数求值顺序等问题）(3)

推荐排行榜

1. 鼠标滚轮实现窗体图形缩放MouseWheel(2)
2. VS中代码对齐等快捷键(2)
3. MySQL中char、varchar和text的区别(1)
4. Python误区之strip,lstrip,rstrip(1)
5. Singleton: this & instance(1)

最新评论

1. Re: 主机名 域名 网站名 URL
mail是主机名这个说法，不知是从哪里来的？ 跟下面两个出处的说法不一致： Examples of host names on the internet could be: – www.cnn.com...
--冷处理
2. Re:Python 函数名作为字典值
谢谢博主，学到了
--熊猫还是猫
3. Re:MySQL中char、varchar和text的区别
thanks
---hello🍒,ukyo🍷~
4. Re:鼠标滚轮实现窗体图形缩放MouseWheel
大佬，Camera类能否提供下？多谢
--ly0329
5. Re:Python 函数名作为字典值
谢谢博主，学习了
--mangosk