

# Shell编程和Wget应用——灌水机

王春昊 (ciclo)

<http://chunhao.net>

June 7, 2009

# Outline

- 1 Shell编程
- 2 wget和灌水机
- 3 Putting It All Together

# 我不是专家

- 我不是Shell编程专家
- 也不是命令行“控”
- 分享经验、乐趣
- 介绍使用电脑、解决问题的全新思维
- 真正成为电脑的主人

# Shell

## Definition (Wikipedia)

**Unix Shell**也叫做命令行界面，它是Unix操作系统下传统的用户和计算机的交互界面。用户直接输入命令来执行各种各样的任务。

## Definition (我的理解)

**Shell编程**就是工具（Linux命令） + 条件 + 控制

# 为什么要写脚本

为什么不用C语言?

一行Shell脚本胜过一万行C代码

```
#include <stdio.h>
#include <stdlib.h>

int main(int argc, char* argv[])
{
    int i = 0;
    ...
    ...
    ...
    if (...) {
        ...
    }
    for (...) {
        ...
    }
    ...
    return 0;
}
```

**head -n 5 input.txt**

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

# 常用的工具

## Don't reinvent the wheel!

- 打印文本——echo
- 显示文件——cat、head、tail
- 随机数——\$RANDOM
- 产生序列——seq
- 计算——expr
- 下载——wget
- .....

- Script Language
  - 脚本语言
  - Script: 剧本
  - 导演: 每个时刻的台词、动作、位置、服装.....
  - 让你的电脑变成舞台, Linux命令是演员, 你是导演

- Script Language
- 脚本语言
- Script: 剧本
- 导演: 每个时刻的台词、动作、位置、服装.....
- 让你的电脑变成舞台, Linux命令是演员, 你是导演

- Script Language
- 脚本语言
- Script: 剧本
  - 导演: 每个时刻的台词、动作、位置、服装.....
  - 让你的电脑变成舞台, Linux命令是演员, 你是导演

- Script Language
- 脚本语言
- Script: 剧本
- 导演: 每个时刻的台词、动作、位置、服装.....
- 让你的电脑变成舞台, Linux命令是演员, 你是导演

- Script Language
- 脚本语言
- Script: 剧本
- 导演: 每个时刻的台词、动作、位置、服装.....
- 让你的电脑变成舞台, Linux命令是演员, 你是导演

# 控制语句

for循环

```
for i in `seq 0 10`  
do  
    something  
done
```

# 控制语句

## while循环

```
while [ $res -eq 0 ]  
do  
    something  
done
```

# 条件语句

## if条件

```
if [ $res -eq 0 ]; then
    dosomething
fi
```

# 更多更详细的语法

- 随用随学
- Google
- Ubuntu中文论坛——Shell脚本版

# Outline

- 1 Shell编程
- 2 wget和灌水机
- 3 Putting It All Together

前端：浏览器、javascript

```
<body>
  <h1>Title</h1>
  <p>content</p>
  <p>content</p>
  <p>content</p>
  <p>content</p>
</body>
```

后台：服务器、php、python

```
<body>
  <h1>Title</h1>
  <? for (i = 0 to 5) { ?>
  <p>content</p>
  <? { ?>
</body>
```

# 灌水机

- 用浏览器发帖，点提交按钮后发生了什么：向服务器发送请求、获得返回页面
- 灌水机：模拟浏览器的行为
- 定制帖子的内容
- 设定发帖的间隔：随机时间

# 灌水机

- 用浏览器发帖，点提交按钮后发生了什么：向服务器发送请求、获得返回页面
- 灌水机：模拟浏览器的行为
- 定制帖子的内容
- 设定发帖的间隔：随机时间

# 灌水机

- 用浏览器发帖，点提交按钮后发生了什么：向服务器发送请求、获得返回页面
- 灌水机：模拟浏览器的行为
- 定制帖子的内容
- 设定发帖的间隔：随机时间

# 灌水机

- 用浏览器发帖，点提交按钮后发生了什么：向服务器发送请求、获得返回页面
- 灌水机：模拟浏览器的行为
- 定制帖子的内容
- 设定发帖的间隔：随机时间

# wget

wget的参数:

```
$wget -help
HTTP options:
...
-referer=URL
-load-cookies=FILE
-post-data=STRING
...
```

# Outline

- 1 Shell编程
- 2 wget和灌水机
- 3 Putting It All Together

# 实战：写一个灌水机

需要工具：

- Firefox插件：Live HTTP headers, Export Cookies
- wget

[http://chunhao.net/blog/  
how-to-write-a-spamming-robot](http://chunhao.net/blog/how-to-write-a-spamming-robot)

# 其他工具推荐

- 处理字符流：sed、awk
- 查找、匹配文件内容：grep
- 查找文件：find
- 定时操作：cron、crontab
- 更复杂的功能：python
- .....

# 不仅仅是灌水机

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 不仅仅是灌压机

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 不仅仅是灌水区

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 不仅仅是灌水机

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 不仅仅是灌压机

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 不仅仅是灌水机

- 监视网页(车票、飞机票、选课网)  
<http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html>
- 定制天气预报  
<http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html>
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
<http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510>
- 自动下载、显示歌词  
<http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420>
- 重复性的工作、体力活

# 不仅仅是灌水机

- 监视网页(车票、飞机票、选课网)  
[http://blog.solrex.cn/articles/  
auto-track-info-with-linux-cli-tools.html](http://blog.solrex.cn/articles/auto-track-info-with-linux-cli-tools.html)
- 定制天气预报  
[http://blog.solrex.cn/articles/  
diy-free-weather-forecast-sms.html](http://blog.solrex.cn/articles/diy-free-weather-forecast-sms.html)
- 批量处理文件  
<http://chunhao.net/blog/command-conquer>
- 暴力破解密码 (并行)
- 游戏 (俄罗斯方块)  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=137510](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=137510)
- 自动下载、显示歌词  
[http://forum.ubuntu.org.cn/viewtopic.php?f=  
21&t=174420](http://forum.ubuntu.org.cn/viewtopic.php?f=21&t=174420)
- 重复性的工作、体力活

# 建议

- 多用man、--help
- 多用Google
- 试着写，尽管刚开始会很慢
- Ubuntu中文论坛——Shell脚本版
- 保留写过的脚本

# 谢谢!

下载此幻灯片:

<http://chunhao.net/pub/zjulug0906.pdf>