



IT 技术从 现在 起 飞



送9264元话费

沃·3G

两年不用交话费新老号均可办理月送1.6G流量、1250分钟话费

搜索

用户名：密码：

飞诺网 加入收藏

用 C 语言编写一个网络蜘蛛

www.diybl.com 时间：2007-12-23 作者：佚名 编辑：andyoo 点击：5465 [评论]

纸张定量取样器 符合国际标准
润湖公司专业生产销纸箱检测设备 服务上门,是您身边纸箱检测设备专家
www.gzrunhu.com

北京总院 治疗间质性肺炎最好
间质性肺炎治疗最好的医院首选北京武总 国家首批公立三甲医院,首批医保定点医院
www.feibing999.com

成功男人晚上都喜欢的"节目"
节目现场直播,午夜频道精彩刺激! 全国送货上门,货到付款
www.3rhdyf.com

```
/*
*****
*filename: mailaddrsearch.c
*purpose: 用 C 语言编写一个网络蜘蛛来搜索网上出现的电子邮件地址
*tidied by: zhoulifa(zhoulifa@163.com) 周立发(http://zhoulifa.bokee.com)
Linux爱好者 Linux知识传播者 SOHO族 开发者 最擅长C语言
*date time:2006-08-31 21:00:00
*Note: 任何人可以任意复制代码并运用这些文档，当然包括你的商业用途
* 但请遵循GPL
*Thanks to: www.gd-linux.org 广东省 Linux 公共服务技术支持中心
*****/

int main(int argc, char ** argv)
{
    int WebPort;
    char * WebHost = 0, * PageAddress = 0, * WebDir = 0;

    if(argc < 2) {if(DEBUG) fprintf(stdout, "Command error, you should input like this:
%s WebPageAddress1 WebPageAddress2 WebPageAddress3 ...", argv[0]); exit(0);}

    NodeHeader = NodeTail = NodeCurr = 0;
    //setlocale(LC_ALL, "zh_CN.gb2312");
    for(i = 1; i < argc; i++)    {
        ret = GetHost(argv, &WebHost, &PageAddress, &WebPort, &WebDir); /* Get web page info */
        if(ret)    {if(DEBUG) fprintf(stdout, "GetHost error from '%s' ", argv); exit(0);}
        AddInitNode(WebHost, PageAddress, WebPort, WebDir); /* add this page to chain */
    }
    free(WebHost); free(PageAddress);free(WebDir);
    if(DEBUG)    {
        fprintf(stdout, " Display.%5d-", FileNumber);
        DisplayNode(NodeHeader); /* display every node */
    }
    HandleInitNode(NodeHeader); /* handle every page */
    return 0;
}
```

频道地图

- C++
 - C++系列教程
 - C++实例
 - C++技术文档
 - C++/C语言函数
 - Mangos
- C
- C++
- Java
- VB
- Delphi
- Foxpro
- 汇编语言
- 游戏开发
- 移动开发
- 软件工程师
- 软工与管理
- VC
- shell编程
- C#

- C++技术文档相关文章
 - 写出几种常用的排序算法以及...
 - using libxml2连载(2 使用简...
 - C语言函数调用约定
 - System函数与脚本的后台执行
 - MFC程序框架剖析（1）
 - 学习vc的经验之谈
 - C和C++编程和学习文档
 - C/C++中的日期和时间
 - CTime,COleDateTime,CStri...
 - The VC++6.0 C++ source co...
 - 我所用过的 VC 技巧 之 WinS...
 - ACE库的生成
 - 托管与非托管
 - 在MFC类中各种类的指针的获...
 - 用动态申请虚拟内存求任意数...
 - libsvm-2.84 源代码分析（二...
 - 对UTF8编码的初步认识！
 - Turbo C 2中的常用快捷键
 - 南京软博会参观点滴
 - const，mutable，volatile意...
 - Turbo C 编程中“Unable to op...
 - 一步一步创建VC2005解决方案
 - MFC在VS.net2005写的程序...
 - 在tc2.0下，进行复制和粘贴

- C++技术文档热门文章
 - win32 API函数大全
 - C++中STL中的map用法详解
 - CString常用方法简介
 - SOCKET编程登峰造极之完成...
 - 如何制作外挂

- 排序算法汇总
- MFC中edit控件的使用
- C++ std::string --- 你可能不...
- CxImage类库使用说明
- C/C++ 笔试、面试题目大汇总

```
/******  
功能：分析网页  
******/
```

```
void AnalyzePage(WEBNODE * node)  
{  
    int fd;  
    int flength = 0;  
    fd = open(node->file, O_RDONLY);  
    if(fd == -1)    goto __AnalyzeDone;  
    flength = lseek(fd, 1, SEEK_END);  
    write(fd, "\0", 1);  
    lseek(fd, 0, SEEK_SET);  
    mapped_mem = mmap(0, flength, PROT_READ, MAP_PRIVATE, fd, 0);  
    GetEmail(mapped_mem);  
    GetLink(mapped_mem);  
    close(fd);  
    munmap(mapped_mem, flength);  
__AnalyzeDone:  
    close(fd);  
    node->IsHandled = 1;  
    remove(node->file);  
}
```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
							18	19	20	21	22					

[提供虚拟机（intel vt）技术支持](#)

如果图片或页面不能正常显示请[点击这里](#)

 发表评论

[【文章投稿】](#) [【收藏此页】](#) [【飞诺社区】](#) [【发表评论】](#) [【关闭】](#)

上一篇：解决：fatal error C1189: #error : include 'stdafx.h' before including this file for PCH 下一篇：linux的sh脚本编程

C++技术文档推荐文章

- | | |
|----------------------------|---------------------|
| 基于COM接口的Shell扩展技术应用 | CString剖析与详解 |
| 罗云彬VxD教程-虚拟机管理器 | 如何在VC++程序中获得其他程序的图标 |
| 用Socket和MSHTML对象模型创建自己的浏览器 | 谷歌浏览器的源码分析(2) |
| 谷歌浏览器的源码分析(11) | 动画窗口的实现和边框绘制 |
| emacs中对C++的.h语法高亮不对的设置 | MFC ActiveX 去掉安全警告 |
| DLL和进程的地址空间 | C++ Builder入门第一例 |
| 一刻钟精通正则表达式 | 基于对话框的应用程序 |
| 嵌入式C++编程准则 | C小技巧 |

文章评论

暂无数据！

请您留言

昵称：

验证码：

[注册会员](#) [会员登录](#)

发表评论

也许你对以下文章感兴趣

- C#基础学习——异步编程篇
- 注意ACE_Pipe的实现
- 提高链表随机访问效率的一种方案
- GDB调试精粹及使用实例
- 在tc2.0下，进行复制和粘贴
- 用VC调试模式不能看到正确地变量的解决办法
- 轻松实现类VC界面（停靠窗口）
- 《C++编码规范》修订说明



飞诺网 鲁ICP备10012328号
IT技术从现在起飞