

基于 P2P 的文件共享系统的设计与实现

范会波,张新有

(西南交通大学 信息科学与技术学院,四川 成都 610031)

摘要:随着网络技术的发展,人们可以从网络中获取越来越多的资源。特别是多媒体技术的发展更是推动了人们对信息共享的需求。对于企业内部来说,用户之间资源共享是必不可少的,而企业员工的计算机存储设备利用率普遍不高,存在一定程度的浪费。文中将 P2P 技术与网络应用中的文件共享相结合,整合空闲的存储资源,实现一个基于 P2P 的文件共享系统。提高了企业用户之间资源共享效率,降低企业的存储成本。目前,P2P 技术越来越成为当前技术的热点,解决了许多传统网络应用中的难题。

关键词:P2P;文件共享;套接字

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2010)03-0048-04

Design and Implementation of File-Sharing System Based on P2P

FAN Hui-bo, ZHANG Xin-you

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: With the rapid development of computer technology, people can acquire more and more resources from the network. Now, the development of multimedia technology which makes the behavior of sharing be developed to a higher level. For the enterprise, employees share resources is essential, the equipments of storage in computer don't be utilized enough. This paper combines the P2P technoes and file-sharing of network application, unite the idle storage resources, implements a file sharing system based on P2P, reduces storage costs, making file sharing easily. At present, P2P become a technical hotspot in the world which resolved many of problems of the traditional web application.

Key words: peer-to-peer; file sharing; socket

0 引言

进入网络时代后,网络规模逐渐增大。网络用户可以方便地获得丰富的信息,用户之间共享资源、互相交流也变得越来越普遍。在这种局势下,文件共享就成为人们获取大量资源的一种有效途径,也是早期 P2P(Peer-to-Peer Network,对等网络)研究和应用的一个主要方向。可以说,正是文件共享的需求激发了 P2P 研究的热潮。在传统的 C/S 结构中实现文件共享需要服务器的参与,客户端将所要共享的文件内容上传到某个特定的服务器,用户就可以到这个服务器上查找、下载自己所需要的文件。所有的资源都存储在服务器上,客户端必须通过服务器连接下载,这就

对服务器的性能提出了很高的要求,使服务器成为整个系统的瓶颈,严重影响了共享资源的用户数目以及下载资源的速度。而 P2P 则改变了 C/S 的网络体系结构,将“中心”的服务器移到了“边缘”的客户机上^[1]。系统中的每一个节点的地位都是平等的,通过节点之间的相互操作,就可以完成文件的共享等相关工作。

该论文就实现了一个企业内部基于 P2P 的文件共享系统,该系统提高了文件检索和共享的效率,同时解决了服务器瓶颈的问题。由于 Java 语言对网络编程的很好支持及其特有的跨平台性,本系统采用了 Java 作为开发语言,采用 Eclipse 作为开发平台。Eclipse 是一个高效的、开源的 Java 开发平台,它的安装包集成了 Java 开发环境的插件 JDT^[2]。

1 典型的 P2P 系统

从 1999 年至今,P2P 系统正逐渐成为研究和应用的热点。P2P 共享系统就是通过网络将端用户设备共

收稿日期:2009-07-02;修回日期:2009-10-21

基金项目:国家自然科学基金(60773102)

作者简介:范会波(1985-),男,山东烟台人,硕士研究生,研究方向为 P2P 网络、P2P 缓存协作;张新有,副教授,研究方向为 P2P 技术、新型网络体系结构、网络管理。

享的资源整合起来,实现大规模的文件共享和存储^[3]。现在已有一些比较著名的系统如: Napster, Gnutella, Kazaa, Morpheus, Chord, BitTorrent 等。以下是两种典型的 P2P 系统:

(1) Napster.

Napster 采用的是部分集中式的节点组织结构。并没有完全取消服务器,只是弱化了服务器的功能^[3]。在 Napster 中,节点在加入系统时将告诉服务器自己机器上所共享的文件,服务器并不存储这些文件的具体内容,而只是保存这些文件的元数据。当一个客户机查找某个文件时,先向服务器提出请求,服务器根据自己保存的元数据查找到文件所在的节点,将该节点的相关信息返回给客户机^[4]。客户机然后直接与文件所在的节点建立连接,下载文件。采用 Napster 除了可以快速定位文件之外,还可以方便地实现服务的商业化。

图 1 表示了 Napster 系统结构图。Client1 先向 Server 发送询问(Query), Server 查询用户请求的资源,将结果(共享文件地址,如 Client2)通过 Reply 返回 Client1,之后 Client1 直接从 Client2 获取所需要的文件。

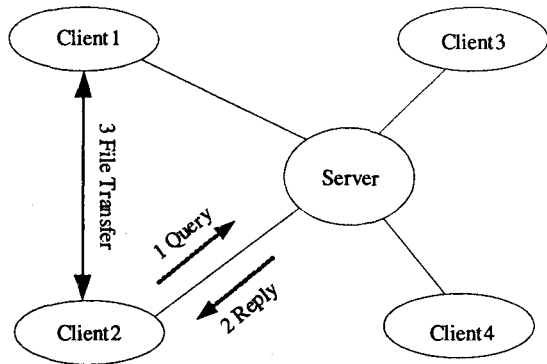


图 1 Napster 系统结构图

(2) Gnutella.

与 Napster 不同, Gnutella 取消了以服务器为核心的目录式结构,而是采用了泛洪算法(Flooding),按宽度优先原则搜索资源的位置^[5]。要查找某个文件,节点首先向自己相邻的活动节点以广播方式发送查询请求包。其他节点在接到包后检查本地是否有符合条件的文件,如果有,则按查询请求包的路径返回一个响应包,该响应包中包括节点的 IP 地址、服务端口号等信息;如果没有,则节点将该查询包发送给它的邻节点,直到查询包的 TTL(Time To Live)计数器减为 0 为止,该计数器中数值是由查询包所发起的节点设置的,以防止该查询无限制地转发下去^[6]。

发起查询的节点收到查询响应包后,则根据包中的信息直接与目标节点建立连接,下载文件。图 2 表示了 Gnutella 查询过程。

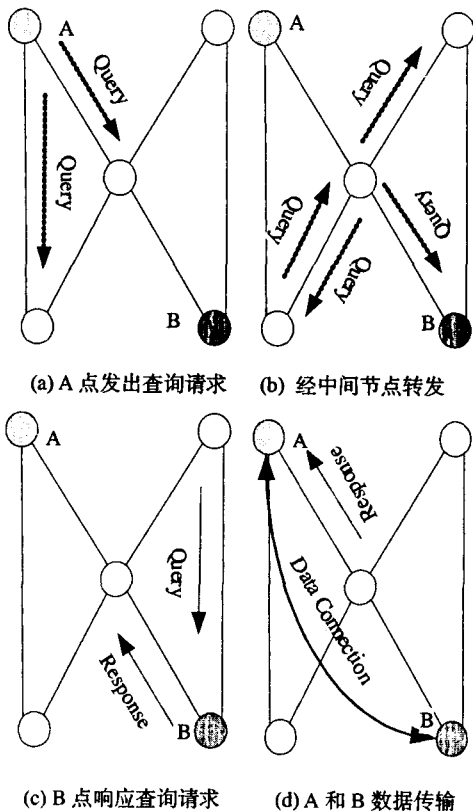


图 2 Gnutella 查询过程

2 系统总体结构及实现流程

2.1 总体结构

本系统主要用于企业网络或园区网内部用户共享资源。采用 P2P 技术保证用户的随时接入、离开不会影响到系统的总体结构。为避免企业的财务、人事等重要文件被不相关人员获知,系统引入了权限管理。用户注册时用户名的开头应为部门的首字母(如研发部为 YF),服务器根据用户的职务和部门赋予用户不同等级的权限,用户在上传文件时必须对文件设置权限并经服务器核实(文件的权限不高于上传文件用户自身的权限),这样保证了只有自身权限大于文件权限的用户方可下载文件。

另外,如果用户上传的文件只能本部门或指定部门人员下载。用户可在上传界面选择指定的部门,客户端会在上传消息的头部加注指定部门的代号。如选择研发部,上传消息头部后会加注“YF”,服务器会限定只有研发部的人员能下载该文件。

为了提高信息的检索效率,系统采用了一个弱化功能的服务器。服务器仅保存当前在线客户端所共享文件的相关属性(如文件名及权限、类型、大小、文件所在的客户端)而不保存文件的具体内容^[7]。同时,服务器可接受客户端发来的反馈信息,对共享文件进行更

新,最大程度保证共享的文件内容是正确的。总体结构如图 3 所示。

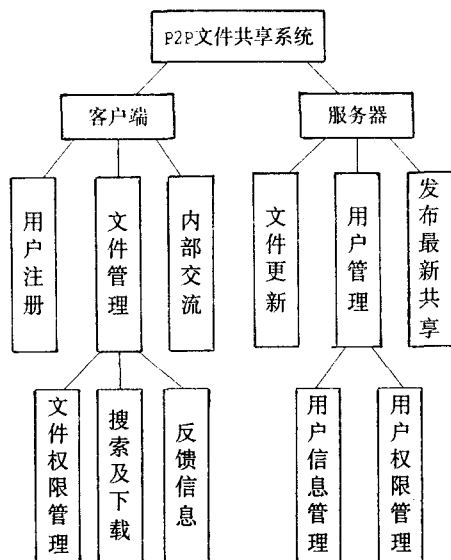


图 3 P2P 文件共享系统结构

由于本系统针对企业内部用户,因此用户注册采用实名制。用户共享的文件经服务器端权限检查通过后,直接显示在服务器的文件列表里,其他满足用户的权限即可通过 P2P 下载。

2.2 系统工作流程

打开系统的客户端软件,登陆服务器便可获得服务器返回的最新共享信息。节点向服务器发出搜索文件请求,服务器将符合该请求的所有节点的 IP,端口号等信息发给该节点。然后节点 Ping 每一个 IP 地址,根据返回时间与最近的节点建立连接,下载文件。图 4 显示了客户端 A 共享文件,客户端 B 搜索、下载文件的工作流程。

3 关键模块的设计与实现

为了方便用户使用软件,本系统采用了可视化的界面设计,用户可以在客户端界面上控制本地文件的共享,分类查看所有的共享文件,了解当前最多、最新的下载文件,搜索下载本身所需要的资源。并可向服务器反馈信息,方便共享文件的更新。其中文件上传、搜索、下载是本系统的核心模块,其功能的实现主要基于 socket 机制。

3.1 文件分块及上传

在本系统中将每一个要共享上传的文件分成大小为 100kB(除最后一块)的文件块,并对每一个文件块按顺序编号。文件块的名称为:原文件名 + .part + 编号。分好的文件块放在和源文件同一目录下。将文件分块是为了在网络传输中更好地利用带宽,提高文件传输的效率^[8]。

关键代码如下:

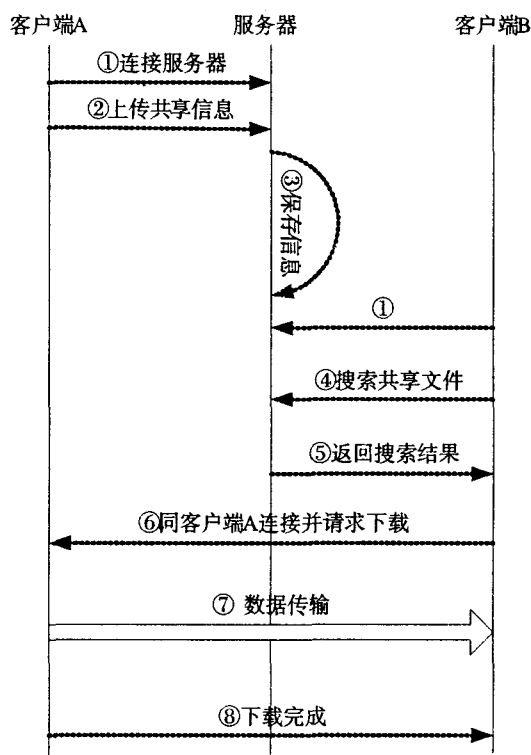


图 4 客户端搜索、下载文件的工作流程

```
Private separatorFile(String sourceFile,long blockSize)
```

```
//根据原文件和文件块大小对文件拆分
```

```
{
    getFileAttribute(sourceFile); //取出原文件的名称及大小
    BlockNumber = getBlockNumber(blockSize); //得到文件块的
    总数
    if(BlockNumber== 1) //如果文件不足 100kB,就一次性写入
    long writeSize= 0; //每次写入的字节
    long writtedSize= 0; //已经写入的字节
    for(int i= 1; i<= BlockNumber; i++)
    {
        if(i< BlockNumber)
        writeSize= blockSize; //取得每一次要写入的文件大小
        else
        writeSize= FileSize- writtedSize;
        CurrentFile= generateSeparatorFileName( sourceFile, i); //生
        成拆分后的文件名
        writeBlockFile( sourceFile, CurrentFile, writeSize, writted-
        Size); //向文件块写入内容
        writtedSize= writtedSize+ writeSize;
    }
}
```

文件拆分成功后,客户端通过 socket 机制同服务器连接,然后将文件的名称及权限、类型、大小、本机的 IP 地址、端口号上传到服务器。为区别于客户端发往服务器的其他消息,上传消息的头部加上一个标识符

(sign="up")。服务器收到消息后从其头部得到标识符,然后利用 writeNodeToLink(LinkList,otherMessage)函数将其他信息写入一个链表中,上传的每个文件都对应着链表中的一个节点。

3.2 文件查看及搜索

本系统中的文件查看和搜索实现的原理是一样的。用户可以在客户端查看网络中所有共享的文件信息,也可查看指定日期共享文件,或者分类别查看共享的文档、软件、代码、书籍等。客户端发出查看信息后,服务器根据存储链表中“类型”关键字查找符合用户要求的文件,并将结果返回给客户端。

本系统的搜索功能是基于关键字搜索,服务器接收到用户要搜索的文件后,调用 searchContent(LinkList,filename)函数到链表中查找是否有匹配的文件,如果找到,则将存储该文件的计算机 IP、端口、文件大小等信息返回给客户端。用户可根据这些信息到相应的计算机上下载所需文件。

3.3 文件的下载与整合

用户搜索到自己所需的文件后利用服务器返回的 IP、端口号同对应的客户端连接。客户端启动 ShareFile 线程输出文件内容。用户则调用 DownFile 类将拆分好的文件块下载到本地保存路径中,同时,用户界面显示下载的进度和速度。

文件下载完毕后,本地保存的只是事先拆分好的文件块,需要进一步将这些文件块整合成一个完整的文件。

文件整合的关键代码如下:

```
private CombineFile()
{
    raf = new RandomAccessFile(sourcefile, "rw");
    for(int i=0; i<FileNum; i++)
    {
        raf.seek(writtedFile);
        fis = new FileInputStream(separatedFilesAndSize[i][0]); //已
        下载的文件块按序号依次将内容写入本地文件
        while((len = fis.read(bt)) > 0)
        {
            raf.write(bt, 0, len);
        }
        fis.close();
        writtedFile = writtedFile + Long.parseLong(separatedFilesAndSize[i][1]);
    }
}
```

3.4 系统运行

使用该系统可以方便企业内部用户交流技术,共享本地计算机的资源。用户不必将企业的数据资源全

部保存到本地,通过本系统可实现对数据的“所得即所需”。并且在企业内部使用时可以达到很高的下载速度,大大节省了用户的时间。

图 5 为软件的下载界面。

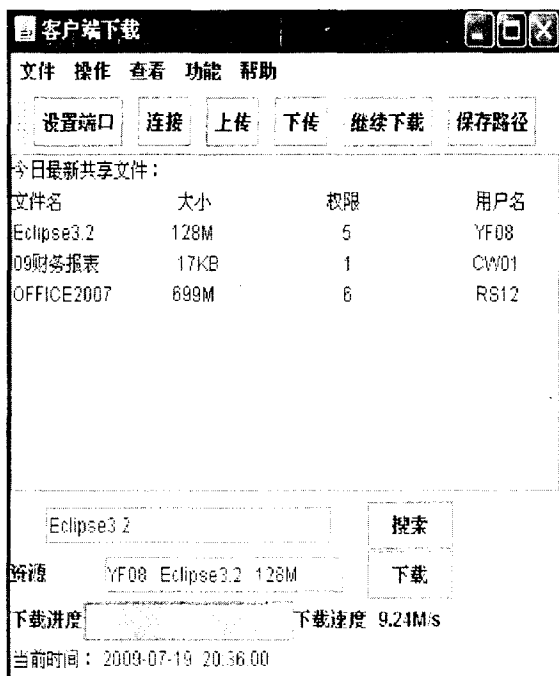


图 5 P2P 文件共享系统下载界面

4 结束语

现代的计算机技术日新月异,企事业内部的计算机性能有了很大的提高,很多计算机闲置了大量的资源。经实际应用证明该系统很好地整合了这些闲置的资源,提高资源的利用率,节省了企业用于数据存储方面的开销。

系统的服务器采用多线程实现,理论上可以无限地接入客户端。并且文件的下传、下载都是在客户端之间进行,有效解决了服务器的瓶颈问题,不会造成网络拥塞。当然该系统对于文件安全还应该有更深入的研究。

参考文献:

- [1] 贾杏丹. P2P 网络搜索技术研究[J]. 计算机应用研究, 2006(3): 31-34.
- [2] Sun Yan, Tang Di. An Algorithm for Curve Clipping Against the Rectangular Window[J]. Computer Applications and Software, 2003, 20(5): 35-36.
- [3] 陈 姝, 方滨兴, 周勇林. P2P 技术的研究与应用[J]. 计算机工程与应用, 2002(13): 25-26.
- [4] 苗 坤, 张 毅, 左 娟. 基于 JXTA 的 P2P 文件共享系统[J]. 计算机系统应用, 2009(2): 37-39.

(下转第 55 页)

