

文章编号: 1005-0523-(2007)04-0093-04

基于C/S架构的eMule监控系统的设计与实现

蔡俊, 黄兆华, 王宝荣

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘要:介绍了P2P和C/S的体系结构,通过修改现有系统eMule,使用三层C/S架构和两层C/S架构,构建了一个eMule监控系统,从而使得改进后的eMule系统的客户端的IP、共享资源、下载完成的资源和是否在线状态受到监控。

关键词:P2P;C/S;eMule监控系统

中图分类号:TP315

文献标识码:A

1 引言

随着网络技术的发展,特别是P2P(Peer-to-Peer)技术的出现,网络上的共享资源逐渐的以指数的形式增长.对于网络用户而言,在网络上找到自己所需要的资源已经是相当容易的事情.但是正是由于共享资源在网络上极大的增长,难免会有不少用户在网络上共享一些不健康的资源.这对网络和社会都是极度不利的.为此,对于网络公司和政府相关部门而言,对P2P软件实行监控,是相当必要的.

本文第二节叙述了P2P,并对当前的P2P系统进行了分类,然后介绍了eMule系统.第三节讲述了C/S体系结构,第四节利用C/S架构在eMule系统基础上构建了一个监控系统,第五节结束语.

2 P2P

P2P是一个分布式的网络系统,由物理上分布的节点组成,这些节点被称为对等实体(Peer).因为系统中分布的每个节点既扮演着服务器的角色,同时也起着客户端的作用,所有节点在系统中的地位都是相同的,它们相互连接,相互协作.

1) P2P系统类型

当前基于P2P的软件是越来越多,而依其搜索架构大致可以分为三类:集中式类型(Centralized)、

广播式类型(Flooding or Gnutella-like)和分散式杂凑表类型(Distributed Hash Table Based, DHT).

(1)集中式类型

在集中式类型的P2P网络系统中,存在一个中央索引服务器(server).当网络中的每个节点要取得网络上其它节点的资源时,都必须向中央索引服务器提出查询请求,然后索引服务器返回查询信息给节点,那么节点就可以直接从接收到的信息中,跟其它节点直接连线下资源,而不再需要中央索引服务器来中转了.最早的Napster^[1]就是使用此种类型最出名的P2P系统.

集中式类型的P2P网络系统虽然效率很高,而且很容易实现,但是却有单点失效和扩展性差的缺点.

(2)广播式类型

广播式类型的P2P网络不再使用中央索引服务器,而是采用泛洪机制进行查询定位.要取得其它节点资源时,由节点向网络中其邻居节点发送搜寻信息,当满足了查询或者超时,将不再继续搜寻;否则,将由其邻居节点继续发送搜寻信息到其邻居节点.Gnutella^[2]是此种类型的最具代表性的系统.

广播式类型的P2P网络系统解决了集中式类型的P2P网络系统的单点失效的缺点.但是同时也导致了网络流量的失控,带宽消耗大等问题.

(3)分散式杂凑表类型

在分散式杂凑表类型的 P2P 系统中, 每一个节点和其提供的资源都要经过杂凑函数(Hash-Function, 如 MD5 和 SHA-1)转换成一串的索引值, 而这一系列的索引值将被放在某一些特点的节点上, 当用户要搜寻某一资源时, 输入的内容也将转换成一串索引值, 通过与分布式杂凑表的对应, 而找到资源的位置. 也就相当于在实际网络上建立了一个结构化的 overlay 层, 即一个逻辑层, 使得更便于信息的查找^[3].

分散式杂凑表类型的 P2P 系统利用较少的路由信息就可以有效地到达目标节点, 而且取消了泛洪算法, 从而节省了带宽, 增强了网络的扩展性.

2) eMule

eMule(电骡), 是以 eDonkey(电驴)协议为基础的文件共享软件, 且是一个开源软件, 在其官方网站以及其它一些网站都能下载到其源代码. eMule 网络是由上百个 eMule 服务器和数以百万的 eMule 客户端组成. 客户端必须连接到服务器来获得网络服务, 这个连接要一直保持到客户端关闭. 服务器之间是没有通讯的, 服务器上提供的是集中的索引服务, 这点类似于 Napster. eMule 是集中式和广播式两种类型的混合型的 P2P 系统^[4].

3 C/S 体系结构

C/S(Client/Server, 客户端/服务器)软件体系结构是由于考虑到资源的不对等(服务器上多, 客户端上少), 以及实现共享两个因素而提出来的. 最初的 C/S 结构(即所谓的两层 C/S 结构)是将应用分为两部分的, 一部分称为前台, 另一部分称为后台. 后台指的是服务器, 只是负责数据的管理, 即服务器上运行的数据库管理系统; 前台指的是客户机, 就是用户通过客户机与服务器上的数据库管理系统取得联系, 对数据库进行需要的操作. 在一般的运用中, 都是数据库与客户端分开的, 服务器上只是存在一个数据库和数据库管理系统, 而客户端上运行着大部分的服务, 比如数据访问规则, 合法性校验等. 每个客户端都是单独与服务器建立连接的.

随着数据库运用的日渐发展, 两层 C/S 结构已经不能满足实际的需求, 因此提出了三层 C/S 体系结构. 三层 C/S 体系结构只是在两层 C/S 结构的基础上添加了一个“中间层”, 也称为组件层或功能层, 其它两层称为数据层和表示层. 表示层位于客户端, 其功能是显示数据于客户端上和发送指令到中间层; 中间层位于单独的服务器上或位于与数据层所

在的服务器上, 其功能是在客户机与服务器之间传输数据; 数据层位于服务器上, 也就是一个数据库管理系统和所需要的数据库. 由于添加了“中间层”, 所以就不需要像两层 C/S 结构那样, 把数据访问规则, 合法性校验等放在客户端, 而是把它们放在中间层. 根据三层 C/S 体系结构中的中间层所在的位置分为以下三种情况(见图 1).



图 1 三层 C/S 体系结构分类

在 P2P 软件三种类型中, 类型 3 很类似两层 C/S 结构, 在实际中是很少使用的, 因为客户端直接与数据库交互, 这样就给数据带来不安全性. 三种类型中, 安全性能最高的是类型 1, 因为组件层与数据层分别位于两台服务器上.

4 系统设计与实现

1) 设计思想

由于 eMule 的使用特别广泛, 客户端数量是以百万为单位的, 两层 C/S 体系结构明显不能满足监控系统的要求, 难以管理大量的客户端^[5], 再考虑到数据安全的缘故, 故采用三层 C/S 体系结构中的类型 1 的模式(见图 1). 这仅是系统的一部分, 只是 eMule 客户端通过中间层服务器与数据库之间数据的交换(见图 2 虚线所示). 另一部分就是监控查询客户端从数据库查询数据, 而监控查询客户端数量很少, 所以没必要采用三层 C/S 结构, 使用两层 C/S 结构就够了(见图 2 实线所示).

系统中数据库服务器采用 MySQL 数据库, linux 操作系统. 中间层服务器中的中间层使用 C 语言编写, 采用多线程的机制. 监控查询客户端采用 VCM-FC 编写, 以图形界面的方式与用户交互.

2) 系统功能

对于监控系统而言, 最关注的就是哪些用户执行了哪些不合要求的操作. 所以 eMule 监控系统的基本功能是监控用户共享的资源, 用户下载的资源, IP 查询和用户数量查询(附加功能).

(1) 共享资源查询: 通过关键词查询网络中所有含有此关键词的共享资源名称和其所在的 eMule 客

户端的 IP 地址,或通过 IP 查询网络中某个客户端所共享的所有的资源名称.

(2)下载资源查询:与共享资源查询类似,也是通过关键词来查询网络中所有含有此关键词的下载了的资源名称和下载该资源的客户端 IP 地址,或通过 IP 查询网络中某个客户端下载了的所有资源名称.

(3)IP 查询:根据 IP 地址查询客户端实际的地址.

(4)用户数量查询:查询当前在线的 eMule 客户端的用户数量和当前断线的 eMule 客户端的用户数量,以及当前连接上 Internet 且开启了 eMule 客户端的用户数量总数(当前连接上 Internet 且开启了 eMule 客户端的用户数量总数=当前在线的 eMule 客户端的用户数量+当前断线的 eMule 客户端的用户数量).

3) 系统数据流程图

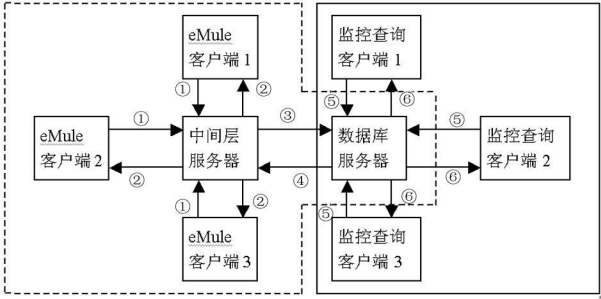


图 2 系统数据流程图

虚线所围区域为三层 C/S 体系结构,实线所围区域为两层 C/S 体系结构.下面是各数据流的含义.

(1) 代表 eMule 客户端向中间层服务器发送的指令.包括:客户端启动后,首次连接上 eMule 服务器时,向中间层服务器先发送其登陆消息,再发送共享列表;客户端下载完某个资源文件后,向中间层服务器发送其资源名称;客户端增加共享资源时,向中间层服务器发送新共享的资源名称;客户端减少共享资源时,向中间服务器发送停止共享的资源名称;客户端与 eMule 服务器断开时,发送下线指令给中间层服务器;客户端退出时,发送退出指令给中间层服务器,把数据库服务器中所对应的内容清除.

(2) 代表中间层服务器向 eMule 返回与①相应的确认信息.发送不成功,要求重新发送.中间层服务器将收到的指令经过处理转换成 SQL 指令发送给数据库服务器.

(3) 代表中间层服务器向数据库服务器发送的数据库操作指令.通常的操作指令是向数据库添加

信息.

(4) 代表数据库服务器向中间层服务器返回的确认信息.操作不成功,要求重新发送.

(5) 代表监控查询客户端向数据库服务器发送的数据库查询指令.

(6) 代表数据库服务器向监控查询客户端返回的数据,在监控查询客户端上显示出来.

4) 数据库设计

监控客户端监控的只是 eMule 客户端的共享资源和下载完成的资源,因此数据库设计为四个字段.

表 1 数据库字段

eMule 客户端 IP	共享资源	下载完成资源	在线标志
--------------	------	--------	------

每一个 eMule 客户端对应于数据库中的一条记录,eMule 客户端 IP 是关键字.数据库中只保留在线客户端和断线而未退出 eMule 的客户端的 IP,共享资源和下载完成资源.eMule 客户端 IP 数据类型为字符串类型;共享资源的数据类型为字符串类型,每个资源名称之间用特定的标志“|”隔开;下载完成资源的数据类型为字符串类型,每个下载完成资源的名称之间也用特定的标志“|”隔开;在线标志数据类型为布尔类型,表示 eMule 客户端是否在线,True 表示在线,False 表示离线.eMule 客户端关闭前,数据库中相应的记录将被删除,以保证数据库大小始终维持在理想的状态下.在线标志表示 eMule 客户端是否在线,通过此字段,可以统计当前连接上 Internet 且开启了 eMule 客户端的用户数量,以及当前在线的 eMule 客户端的数量和当前断线的 eMule 客户端的数量.

5) 中间层服务器实现的关键技术

中间层服务器与 eMule 客户端的连接采用“面向连接”的 TCP(Transmission Control Protocol,传输控制协议)和“面向非连接”的 UDP(User Data Protocol,用户数据报协议)混合的方式,中间层服务器与数据库服务器采用 TCP 连接.其中,为了保证数据的可靠性,eMule 客户端在启动之后,首次连接上 eMule 服务器时,使用 TCP 与中间层服务器通信;eMule 客户端关闭时,使用 TCP 与中间层服务器通信;考虑到提高被监控的 eMule 客户端的性能,eMule 客户端和中间层服务器在其它情况下的通信都采用“面向非连接”的 UDP.eMule 客户端向中间层服务器发送共享列表采用消息递增的方式,即,有几个共享文件,就发送几个 UDP 包.中间层服务器接收到客户端的数据之后,对数据进行解析,将其转换成相应的 SQL 语句,对数据库服务器进行相应的操作.

中间层数据连接图如下图(图 3)所示. “[IP, Share, Downloaded, State]”中的 IP 表示 eMule 客户端的 IP 地址, Share 表示共享资源的名称, Downloaded

表示下载完成的资源名称, State 表示在线与否, True 表示在线, False 表示离线. “TCP”和“UDP”分别表示连接方式是 TCP 和 UDP.

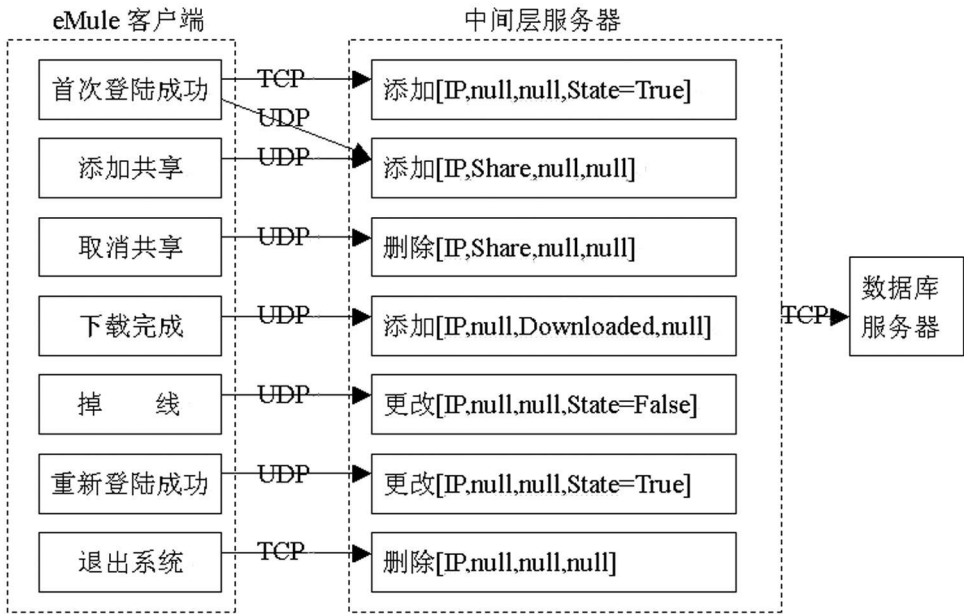


图 3 中间层数据连接图

5 结束语

目前 P2P 资源共享软件中的资源尚且没有得到有效的控制,很多不健康的资源在 P2P 网络上肆意的传播.本系统的研究和开发有利于公司以及政府相关部门对 eMule 的用户进行监控,在一定程度上对 eMule 网络进行一定的规范.同时,也为其它 P2P 资源共享软件的监控系统的设计起到了一个范例的作用.

参考文献:

[1] Napster. <http://www.napster.com/>
[2] The Gnutella Protocol Specification v0.4 Revision 1.2
[3] 周健,洪佩琳,李津生. DHT 网络中一种基于树型结构的负载均衡方案[J]. 小型微型计算机系统, 2006, 27(11): 2042—2046.
[4] 俞燕燕,李绍滋. eMule 系统协议分析[J]. 合肥工业大学学报, 2006, 29(9): 1096—1100.
[5] 倪晓明,付碧丹. 三层 C/S 结构及其应用开发[J]. 四川通信技术, 2001, (31): 9—11.

Design and Implementation of eMule’s Monitoring System
Based on the Structure of C/S

CAI Jun, HUANG Zhao-hua, WANG Bao-rong

(School of Infumation Eng., East China Jiaotong University, Nanchang 330013, China)

Abstract: This paper presents P2P and the structure of C/S. The paper makes use of structures of two-tier C/S and three-tier C/S through revision of the existing system eMule to build a eMule’s monitoring system in order to make IP, shared files, downloaded files and the state of online of the client of the revised system of eMule monitored.

Key words: P2P; C/S; eMule’s monitoring system