一种基于P2P机制的虚拟以太网改进设计与实现

李立新[1] 周雁舟[1] 常光辉[2] 李文俊[1]
(解放军信息工程大学电子技术学院 郑州 450004)[1] (重庆大学计算机学院 重庆 400044)[2]

摘 要 传统基于C/S模式的虚拟以太网的数据在处理过程中需要两次经过协议栈,这在一定程度上影响了远程安全组网的传输效率。提出了一种改进的基于P2P机制的虚拟以太网设计,实现虚拟网络中节点之间的对等连接,能有效克服传统虚拟以太网的不足。测试表明,该设计提高了系统的通信效率,明显改善了虚拟以太网的传输性能。

关键词 虚拟以太网,隧道,P2P
中图法分类号 TP393.08 **文献标识码** A

P2P Based Improved Design and Implementation of Virtual Ethernet

LI Li-xin[1] ZHOU Yan-zhou[1] CHANG Guang-hui[2] LI Wen-jun[1]
(Institute of Electronic Technology, PLA Information Engineering University, Zhengzhou 450004, China)[1]
(College of Computer Science and Technology, Chongqing University, Chongqing 400044, China)[2]

Abstract The data processing of traditional C/S mode based virtual Ethernet enters the protocol stack twice, which affects the transmission efficiency of long-distance safety networking to a certain extent. This paper presented an improved virtual Ethernet design based on P2P mechanism, it realizes P2P-based connection among the virtual Ethernet network nodes, which efficiently overcomes such problem of traditional virtual Ethernet. Test results demonstrate that, besides improvement on the efficiency of the system's communication, the design obviously improves transmission performance of the virtual Ethernet.

Keywords Virtual Ethernet, Tunnel, P2P

由于现有网络的开放性,实现安全的远程组网成为许多涉及敏感信息传输的单位或组织亟待解决的问题。利用虚拟网卡构建虚拟以太网是解决开放公共网络上远程安全组网的一种有效方法,但是传统的虚拟以太网组网方式由于其数据通信机制限制,传输效率相对较低,本文在传统虚拟以太网中引入P2P机制,从而有效地改善了虚拟以太网的通信效率。

1 虚拟以太网的基本概念

1.1 虚拟以太网原理

虚拟以太网技术是对以太网的一种模拟。以太网是一种基于总线的广播网,采用CSMA/CD(带冲突检测的载波侦听多路接入)的媒体接入方法,它的数据封装格式在RFC 894中定义。虚拟以太网技术就是通过将应用数据最终封装成以太网帧,再通过其它方式实现对封装以太网帧的传送。这样对于以太网帧所封装的应用数据而言是工作在一个虚拟的以太网环境之中的,因此称这种技术为虚拟以太网技术。

虚拟以太网技术最重要的任务是仿真出以太网的环境,一方面需要实现和IP协议栈的接口,从而可以对IP协议栈数据进行以太网帧封装相关处理;另一方面由于IP协议栈工作于系统的核心态,封装后的以太网帧需要一个和用户态通信的接口,从而实现封装的以太网帧通过用户态处理后以某种方式进行传输的目的。虚拟以太网的这一功能主要通过虚拟网卡来实现。

1.2 虚拟网卡

虚拟网卡实际上是底层的一个驱动程序,它对用户表现出真实网卡的所有功能,用户可以看到虚拟网卡的连通状态、设置虚拟网卡的IP地址。虚拟网卡在功能上一方面实现和操作系统核心态的TCP/IP协议栈交互,以实现对进出协议栈数据包进行以太网格式的封装和解封装处理,另一方面向用户态的应用层提供访问接口,供数据包在用户态和核心态之间进行传递。

1.3 虚拟以太网数据处理流程

虚拟网卡向用户提供了一个逻辑的网络地址空间,用户的应用数据处在一个虚拟的以太网环境之中,但数据最终还是通过封装后经真实网卡或其它物理链路进行传输。图1显示了虚拟以太网下的数据处理流程。

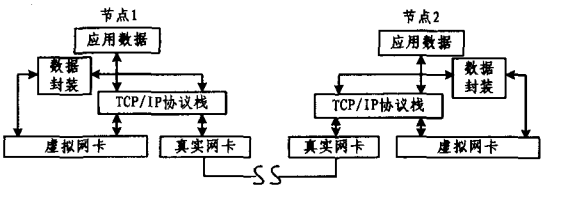

图1 虚拟以太网数据处理流程

到稿日期:2008-05-05

李立新(1966—),男,副教授,主要研究方向为计算机网络、数字证书应用、信息安全,E-mail:PICRIC@139.com;周雁舟 男,副研究员,主要研究方向为可信计算、计算机网络;常光辉 男,博士研究生,主要研究方向为可信计算、分布式计算;李文俊 男,硕士研究生,主要研究方向为信息安全、计算机网络。

节点1需要和节点2进行数据通信时，节点1的应用数据首先发往TCP/IP协议栈，经协议栈封装后由虚拟网卡截获，虚拟网卡将协议数据进一步封装为以太网帧交付应用层，以太网帧在应用层封装后再经TCP/IP协议栈、真实网卡或其它物理链路发送出去，在接收端的节点2中，数据采用和发送端相反的处理流程。这样对于应用数据而言是工作在虚拟网卡所构建成的虚拟以太网之中。

1.4 虚拟以太网实现远程安全组网的性能特点

虚拟以太网将整个以太网帧作为应用层隧道的负载通过其它方式在网络中传输，没有修改协议栈数据包的内容，可以顺利地穿透NAT设备，防火墙只要开放特定的一个通信端口，就可以顺利实现连接，因此，虚拟以太网具有良好的网络适应性。虚拟以太网不需要和特定的应用协议结合，因而具有良好的应用适应性。虚拟以太网的不足在于其数据在处理过程中需要两次经过协议栈，这在一定程度上会影响远程安全组网的效率。

2 典型虚拟以太网软件(OPENVPN VTUN)分析

典型的虚拟以太网软件有VTUN和OPENVPN，这里以功能比较完善的OPENVPN[1]为例进行分析。OPENVPN是一个开源软件，其利用开源虚拟网卡驱动TUN/TAP具有灵活的配置选项，可以根据不同的安全需求在各种复杂网络环境下实现远程安全组网。OPENVPN采用C/S结构，服务器要求具有公网地址，可以供客户端进行连接，客户端地址没有限制，客户端可以顺利穿透一般网络设备和服务器建立安全隧道。OPENVPN根据虚拟网卡驱动TUN/TAP[2]的两种工作模式，可以分别工作于TUN模式和TAP模式，TUN模式是一种虚拟的点到点网络设备，它实现了IP隧道的功能，处理的对象是IP数据包；TAP模式是一种虚拟的以太网设备，它实现的是以太网帧隧道的功能，处理的对象是以太网帧。TUN模式用于网络间构建安全隧道，TAP模式用于将服务器和各个客户端组建成一个虚拟的局域网。对于TAP模式，其数据的流向如图2所示。

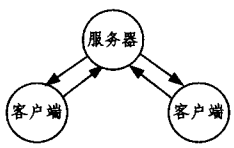

图2 OPENVPN节点间数据传送流程

客户端和服务器之间的通信直接在它们之间的安全隧道中传输，如果客户端之间需要进行通信，它们的通信则经服务器进行转发。由于客户端所处的环境不固定，不一定具有公网地址，它们之间不能确保能直接建立连接，但是服务器具有公网地址，每一客户端都可以和服务器建立安全隧道，通过服务器对客户端之间通信的转发，可以保证客户端之间安全通信的可行性。但是这种方式也存在着一定的问题。当客户端之间存在大量通信时，它们的通信都需要经过服务器进行转发，这样服务器的转发速率将成为整个系统的瓶颈，而且虚拟以太网方式下数据需要两次经过协议栈，隧道的处理本身又需要一定的开销，因此，通过服务器进行转发，其效率会随着客户端通信量的增加而急剧降低，甚至导致系统的崩溃。

以上分析可以看出，虚拟以太网在通信方式上存在着不足，需要对其进行进一步改进。

3 基于P2P的改进设计

基于P2P网络模式的应用技术近几年来发展迅速，P2P(Peer-To-Peer)[3]称为点对点网络或对等网络，它与传统的C/S,B/S模式的通信技术不同，所有节点的地位是平等的，节点间可以直接进行通信，从而摆脱了对服务器的依赖，并且可以克服服务器带来的单点失效及瓶颈问题，为资源共享、即时通信等应用提供了一种高效的机制。传统虚拟以太网中，服务器对客户端通信数据进行转发相当于也是服务器提供的一种服务，也是一种C/S模式的应用，正如第2节中的分析，这种C/S模式的效率很低，因此如果在客户端的通信中引入P2P机制，将能有效克服传统虚拟以太网这种C/S模式的通信机制，提高系统的通信效率，明显改善虚拟以太网的传输性能。根据以上分析，本文对虚拟以太网进行了改进，在虚拟以太网的节点间通信中引入P2P机制，设计了基于P2P机制的虚拟以太网体系结构。

3.1 基于P2P的虚拟以太网节点定位方式

基于P2P的虚拟以太网要解决的问题是如何实现虚拟网络中节点之间的对等连接。要实现节点之间的对等连接首先要解决的问题就是节点定位。P2P网络中节点和资源的定位方式有中心服务器搜索方式、洪泛搜索方式、基于DHT(Distributed Hash Table)搜索方式、混合式P2P网络搜索方式4种。洪泛式搜索方式网络可控性较差，且不易扩展；基于DHT方式和混合式搜索方式实现机制复杂，网络状态维护困难，可以用于资源共享的P2P网络应用中，但不适用于虚拟以太网中节点之间的通信。另外，采用哪种节点定位方式很重要的一点取决于网络拓扑结构，虚拟以太网拓扑中，多个客户端连接到服务器，因此可以让虚拟以太网的服务器在实现对客户端认证和注册的同时兼作中心索引服务器，采用中心服务器搜索的方式，提供节点位置索引查询的服务。

3.2 基于P2P的通信模式设计

根据以上分析，基于P2P的虚拟以太网采用如图3所示的集中式网络结构。

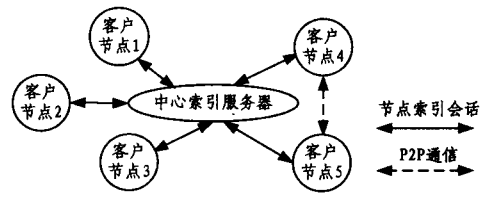

图3 基于P2P的虚拟以太网网络结构

其中客户节点是虚拟以太网系统的基本组成元素，所有客户节点在安全通信时是地位平等的对等实体；虚拟以太网服务器作为中心索引服务器，当客户节点成功登录到虚拟以太网系统后，就可以从服务器获得其它在线节点的地址信息，发起和它们的点对点的通信连接。如果连接建立成功，节点间就可以建立安全的P2P隧道，如果不能建立P2P连接，则它们之间的通信由服务器转发。

客户节点和虚拟以太网服务器通过一个数字证书ID号作为标识。服务器通过CA颁发的数字证书对节点进行认证和注册，登记节点的必要信息，包括虚拟IP地址、通信IP地

址、节点证书 ID 等。系统提供基于 P2P 的节点状态处理和安全通信两种机制。节点状态处理具有灵活的组网功能，包括节点的动态登录和退出、节点的发现与定位；同时通过维护隧道列表和用户列表实现对客户节点的实时维护和状态管理功能。安全通信机制保证了信息在客户节点间的安全传输。

3.3 基于 P2P 的网络节点状态处理

当中心索引服务器初始化完成并正常启动后，VPN[4,5] 节点就可以登录 VPN 中心索引服务器，系统的状态处理过程如图 4 所示。

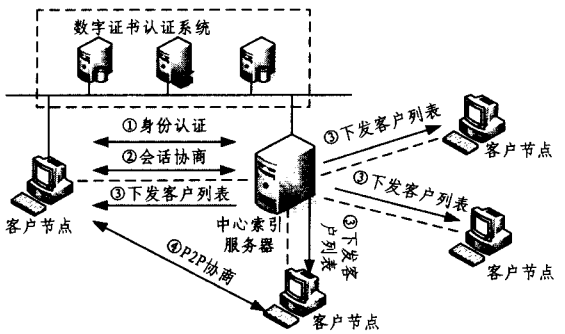

图 4 基于 P2P 的虚拟以太网状态处理过程

3.3.1 节点登录系统

登录的具体过程：

(1)客户节点通过自己的数字证书向中心索引服务器(即虚拟以太网服务器)发起登录认证请求，中心索引服务器通过数字证书认证系统与请求的节点间进行基于数字证书的双向身份认证。

(2)身份验证通过后两者协商建立安全会话，获得 VPN 节点的真实通信 IP 地址。协商的安全会话用来保护客户节点的搜索会话和管理会话，当客户节点间不能建立 P2P 连接时，它们之间的通信也是通过客户节点与服务器协商的安全会话以接力的方式进行保护的。在该安全会话的保护下，新节点上传相关节点信息。

(3)客户节点信息上传完毕后，中心索引服务器更新客户列表和隧道列表，并将客户列表广播给所有客户节点，由客户节点更新自己的客户列表。

至此客户节点成功登录虚拟以太网系统，登录后，所有登录的节点间就可以进行点对点的会话连接。中心索引服务器会定时向客户节点发送心跳检测信息，检测客户节点的状态。

3.3.2 客户节点地址定位

由于公共网络的复杂性，系统客户节点前可能存在 NAT[6](Network Address Translator)，NAT 分为完全圆锥型 NAT(FuUCone NAT)、地址限制圆锥型 NAT(Address Restricted Cone NAT)、端口限制圆锥型 NAT(Port Restricted Cone NAT)、对称型 NAT(Symmetric NAT)4 种类型，对于后 3 种 NAT 之后的客户节点，它们之间建立 P2P 连接相对困难，因此系统中 P2P 连接建立时采用一种简单的定位方式，不考虑网络中的 NAT 问题，服务器对客户节点地址的定位方式采取将客户节点的通信公网地址作为其地址保存在用户列表之中。对于完全圆锥型 NAT 后的客户节点，由于通信公网地址和私网地址一一对应，节点间可以建立 P2P 连接，对于存在复杂 NAT 不能建立点对点连接的情况，节点间的通信由服务器转发。

3.3.3 P2P 安全连接的建立

对于每一客户节点，既可以主动发起和其它节点的点对点连接，也可以响应其它节点的连接请求建立点对点连接。当某一客户节点需要和其它客户节点建立 P2P 连接时，从隧道列表中查找是否已经与该目的客户节点建立安全隧道，如没有则从客户列表中查找该客户节点的详细信息，发起和该客户节点的连接请求，安全隧道的建立过程和节点登录服务器过程相同。由于客户节点间可能存在无法穿透的 NAT 或其它网络设备，如果发起连接请求后固定时间内不能收到对应客户节点的应答则认为无法和该客户节点建立 P2P 连接，此时调用用户管理模块将该用户项的通信地址改为服务器地址，这种情况下客户节点之间的通信由服务器进行转发。

对于每一客户节点，在登录系统后都会和服务器一样启动守护进程，等待其它客户节点的连接。当有客户节点发起和它的连接时，和服务器响应客户节点的连接一样，和该客户节点进行认证并协商安全隧道。

3.3.4 节点退出系统

系统客户节点的退出分两种情况，一种是正常退出，这种情况客户节点向中心索引服务器发送退出消息，从隧道列表中删除与中心服务器的隧道项，中心索引服务器接收到退出消息后作出退出处理；另一种是非正常退出，当中心索引服务器超时不能收到客户节点的心跳检测信息的应答则认为客户节点非正常退出，此时执行客户节点退出处理。

中心索引服务器对节点退出处理：当中心索引服务器确定某一客户节点退出系统后，删除与该节点的安全隧道，清除隧道列表和客户列表中对应的项。同时向其它在线节点发送该客户节点退出消息，其它客户节点收到消息后从自己的客户列表中删除该节点项。

3.4 基于 P2P 的虚拟以太网实现

基于 P2P 的虚拟以太网系统的实现分服务器和客户端两部分。

服务器用于管理整个系统中客户节点的状态，辅助建立客户节点间的 P2P 连接，其主要工作包括：通过心跳信息检测在线客户节点；处理客户节点的登录和退出；在用户列表更新后广播用户列表；响应客户节点的客户列表请求；转发不能建立 P2P 连接的客户节点之间的数据。

其中响应客户列表请求时只需要将客户列表通过和请求客户节点建立的隧道发送给该节点。转发客户节点间数据包时将数据包还原后根据通过与目的节点的隧道发送到相应节点。

VPN 客户节点的 P2P 模块用于获取系统中在线 VPN 节点状态，发起和接受其它 VPN 节点的 P2P 连接，其主要工作包括：应答服务器心跳检测信息，确认节点在线状态；发起和其它 VPN 节点的 P2P 连接；接受其它 VPN 节点的 P2P 连接；接受服务器广播的用户列表；向服务器请求用户列表。

其中应答服务器心跳检测信息时只需要发送确认数据报文，向服务器请求用户列表时向服务器发送请求报文。

4 系统测试

本文在对基于 P2P 的虚拟以太网系统实现的基础上，对系统的 P2P 组网性能和效率进行了简单的测试。系统的测

试平台利用现有局域网段搭建。现有局域网段带宽 100M，网络地址是 25.20.186.0，子网掩码 255.255.255.0，现通过一台试验机器构建 25.20.187.0 网段，子网掩码 255.255.255.0。通过一台双网卡机器作为网关实现两个网段通信的路由，模拟公共网络环境。测试的网络结构如图 5 所示。

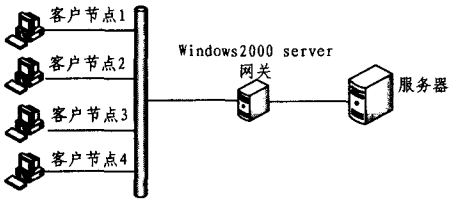

图 5　测试环境图

4.1　P2P 连通性测试

将 Windows 2000 server 网关配置为路由器工作模式，在客户节点 1 建立 FTP 服务器，在客户节点 4 下载客户节点 1 上的数据，使用 tcpdump 在客户节点 4 查看通过真实网卡接收数据的地址，通过检测，可以看到接收的数据包的源 IP 地址是客户节点 1，从而验证了 P2P 连接的有效性。

将 Windows 2000 server 网关配置为 NAT 网关，采用和上面相同的方法查看客户节点 4 基于真实地址的数据接收来源，通过检测，可以看到客户节点 4 接收的数据包的源 IP 地址是 VPN 服务器，而没有接收到客户节点 1 的数据。这一结果是因为客户节点 1 和客户节点 4 之间存在双边 NAT，它们之间无法建立 P2P 连接，通信需要通过服务器转发。

通过以上测试，验证了系统在可建立 P2P 连接时客户节点间 P2P 连接建立的有效性，以及客户节点间不能建立 P2P 连接时通过服务器转发数据的有效性。

4.2　传输性能测试

通过比较 P2P 连接时的系统传输速率和采用服务器转发时的系统传输速率，检测系统 P2P 模式的传输性能以及 P2P 模式对系统效率提高的程度。

在 Windows 2000 server 网关配置为路由器的情况下，客户节点之间可以建立 P2P 连接，系统的 P2P 通信效率在该环境下测试。首先，在客户节点 1 上建立 FTP 服务器，在客户节点 1 与客户节点 4 之间建立 P2P 连接，测试客户节点 4 从客户节点 1 下载数据的速率。其次，在客户节点 1 和客户节点 2 上建立 FTP 服务器，建立客户节点 1 与客户节点 4 的 P2P 连接、客户节点 2 与客户节点 3 的 P2P 连接，客户节点 4 和客户节点 3 同时从客户节点 1 和客户节点 2 下载数据，测试客户节点 4 的数据下载速率，以上两组数据作为 P2P 模式传输速率的测试结果。

为了比较 P2P 模式与服务器转发模式传输性能的差别，对服务器转发模式的传输速率也进行了测试。在图 5 的网络配置下，客户端之间的通信通过服务器转发，服务器转发模式的传输速率测试在该网络配置下进行。测试时采用和上面类似的方法，分别测试单用户传输和多用户传输时系统的传输速率。测试数据如表 1 所列。

表 1　传输速率比较

测试内容 / 序号	直接传输	单用户服务器转发	多用户服务器转发	单用户 P2P 方式	多用户 P2P 方式
1	3582.93	290.78	264.62	300.76	302.45
2	3720.93	285.29	286.10	288.48	297.23
3	3720.93	279.08	265.48	301.05	311.28
4	3784.46	275.46	284.45	311.22	334.29
5	3764.71	339.51	293.85	330.11	330.56
6	3501.81	278.78	253.12	338.76	331.73
7	3688.76	295.44	245.38	312.34	321.57
8	3397.25	287.87	256.52	322.56	330.34
9	3792.59	256.78	264.01	312.45	310.21
10	3818.33	284.36	278.60	323.34	308.31
平均值	3677.27	287.34	269.21	314.10	317.80

通过比较系统传输速率，单用户传输时，系统的 P2P 方式的系统传输速率相对服务器转发方式传输速率提高了 9.31%，这是因为服务器转发方式下数据需要经过服务器方的中转处理，服务器对数据包的转发处理需要一定的时间开销，而 P2P 方式是直接进行数据的传输，效率相对较高；在多用户传输的情况下，P2P 方式的系统传输速率和单用户传输情况下的速率基本保持一致，而相对服务器转发方式多用户传输速率提高了 18.05%，以上数据也表明，客户端之间的通信关系增加时对 P2P 方式的系统传输效率影响很小。另一方面，服务器转发方式下多用户传输速率相对单用户传输速率降低了 10.6%，可以看出，随着客户端之间通信关系的增加，服务器转发方式的系统传输速率会逐渐减小。

以上测试结果及分析可以看出，P2P 的通信方式相对服务器转发方式在系统传输速率方面具有明显的改善作用。

结束语　虚拟以太网可以实现具有良好灵活性和适应性的远程组网功能。但传统虚拟以太网存在传输效率较低的不足，本文针对这一不足在虚拟以太网中客户节点间的通信中引入了 P2P 机制，用于改善系统的传输效率。测试证明，在引入 P2P 机制后，系统的传输速率有了明显的改善，相对于传统虚拟以太网，引入 P2P 机制后系统的效率受客户节点数量的影响相对很小，对于将虚拟以太网应用于较多用户节点的应用环境具有较高的应用价值。

参 考 文 献

[1] Hosner C. Open VPN and the SSL VPN Revolution HTTP://openvpn.sourceforge.net/resource/sslvpn.pdf

[2] 麻利辉. 虚拟网卡 TUN/TAP 驱动程序设计原理. http://www-900.ibm.com/developerworks/cn/linux/l-tuntap/index.shtml

[3] 杨天路，刘宇宏，等. P2P 网络技术原理与系统开发案例[M]. 北京：人民邮电出版社，2007，6：84

[4] Younglove R. Virtual private networks-how the work. Computing & control Engineering Journal，2000，11(6)：260-262

[5] 戴宗坤，唐三平. VPN 与网络安全[M]. 北京：电子工业出版社，2002：1-10

[6] Aboba B，Dixon W. IPsec-Network Address Translation (NAT) Compatibility Requirements. RFC3715. March 2004