

基于云计算的医学图像流媒体应用研究

赵美泽¹ 刁丽娟²

(鲁东大学 烟台 264025) (华东师范大学信息科学技术学院 上海 200241)²

摘要 提出了基于云计算的医学图像流媒体服务传输系统,此系统将医疗设备和 PACS 进行整合,将来自于各个医疗设备的图像用扫描仪扫描为数码图像,并以流媒体的形式实时提交给远程服务器端。在云数据库虚拟化存储平台上,用先进的流媒体技术解决了医疗资源共享和高分辨率图像连续存储问题。实验结果表明,与传统服务器在 LAN (Local Area Network)、TANet(Taiwan Academic Network)和家庭网络中使用 FTP 传输相比,建立的云计算医学图像流媒体服务传输系统的网络传输率分别提高了 45.6%、49.4%和 8.1%。

关键词 云计算,流媒体,图像归档和通讯系统,医学数字图像通讯,虚拟化,Xen 服务器

中图分类号 TP391 **文献标识码** B

Research of Applying Streaming Media for Medical Image Based on Cloud Computing

ZHAO Mei-ze¹ DIAO Li-juan²

(Ludong University, Yantai 264025, China)¹

(Department of Computer and Technology, East China Normal University, Shanghai 200241, China)²

Abstract This paper proposed the medical image streaming media transmission system based on cloud computing, which integrates medical devices and PACS. The system makes use of scanning to transform images from different medical into digital images which are submitted to the remote servers real time in the form of streaming media. In the virtual storage platform of cloud database, the paper adopted advanced streaming media technology to solve the problems of sharing medical resource and storing continuously high resolution images. The experimental results show that compared with the transmission rates of traditional servers using FTP in LAN (Local Area Network), TANet (Taiwan Academic Network) and the home network, the transmission rate of the medical image streaming media transmission system based on cloud computing rises respectively 45.6%, 49.4%, 8.1%.

Keywords Cloud computing, Streaming media, PACS, DICOM, Virtualization, Xen server

1 引言

伴随信息和通信技术的发展,在医院里病人的检验报告和通过医疗设备拍的体检图片等资料都被处理成为电子信息资源,虽然电子信息资源比纸质的信息用起来更方便、更减少时间,但是解决医疗信息资源的存储和共享是一个关键的问题。为了解决数据储存和软件不兼容性问题,提高图像数据传输率,我们提出了基于云计算的医学图像流媒体服务传输系统。此系统通过互联网可以与 PACS 协调工作,并将大量 DICOM(医学数字图像)文件准确无误地传输到云数据库虚拟化存储平台上。此系统主要解决了如下问题:①解决了以高传输率传输大容量图像文件的问题;②解决了传输大批量图像文件的稳定性问题;③解决了医疗资源存储和共享问题,即医生或病人可在任何地点任何时间及时浏览和上传病人病历的电子信息,这给医生和病人带来了很大的便利,避免了重复体检的人力和物力。

2 技术基础

2.1 云计算和流媒体技术

云计算是一个基于分布式计算的新技术。云计算的一个显著特点是它集成了不同的平台、架构和规模的计算机。在服务器和客户端之间有 3 层^[1],如图 1 所示。应用层是指在云应用程序体系结构的用户可以不安装其他程序或软件就可完成工作,从而解决了软件维护的负载;平台层提供了一个云基础设施平台,它可以减少部署软件、硬件的成本,并使管理过程简单;基础设施层应对交付和计算,它通常是一个虚拟化过程。

云计算包含了 3 层含义^[2]:用户的公共性、设备的多样性、商业模式的服务性。

根据云的部署和开放程度的不同,云可分 4 类^[3]:公有云 (Public Cloud)、社区云 (Community Cloud)、私有云 (Private Cloud)和混合云 (Hybrid Cloud)。

到稿日期:2012-09-13 返修日期:2012-12-13 本文受上海市国际科技合作基金项目(11530700300)资助。

赵美泽(1961-),女,硕士,高级实验师,主要研究领域为云计算、图像学,E-mail:ytwflw@126.com;刁丽娟(1981-),女,博士,主要研究领域为云计算、图像处理、语义计算。

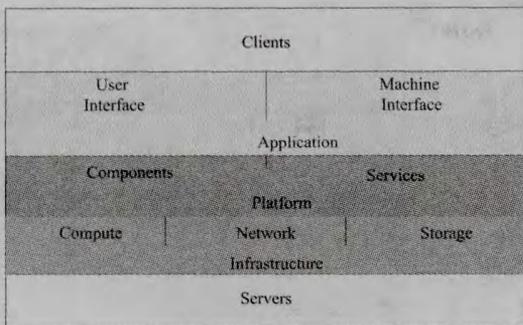


图1 云计算的层次模型

流媒体(Streaming Media)实现的关键技术就是流式传输,流式传输主要指将整个音频和视频及三维媒体等多媒体文件经过特定的压缩方式解析成一个个压缩包,由视频服务器向用户计算机顺序或实时传送^[4]。在流媒体传输中,标准的协议就是 RTP(Real time Transport Protocol,实时传输协议)、RTCP(Real-time Transport Control Protocol,实时传输

控制协议)、RTSP(Real Time Streaming Protocol,实时流媒体协议)和 RSVP(Resource Reserve Protocol,资源预订协议)^[5]。

2.2 图像归档通讯系统(PACS)和医学数字图像通讯(DICOM)

PACS系统是Picture Archiving and Communication Systems的缩写,意为影像归档通信系统。它是应用在医院影像科室的系统^[6]。医学数字图像通讯DICOM3.0(DICOM, Digital Imaging and Communications in Medicine)是医学影像信息学领域的国际通用标准^[7]。

3 医学图像流媒体云服务系统设计

3.1 系统整体架构

我们的系统整体构架如图2所示,系统主要包括两大部分:第一部分是前端工作站,第二部分是云服务平台,其中:云服务平台由云代理服务器和后端云数据库组成。

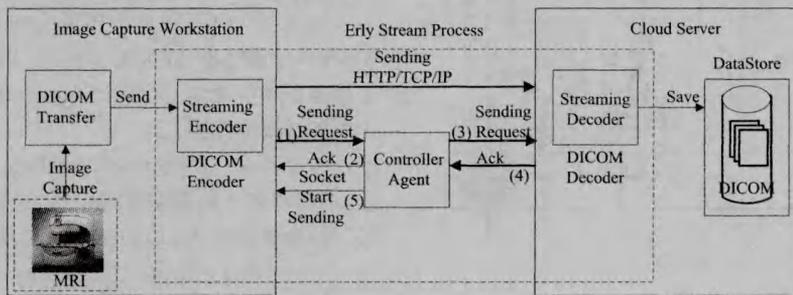


图2 系统整体构架

前端工作站功能是通过MRI获取医疗图像,并将获取的医疗图像转化为DICOM格式存储。工作站当接收到云代理服务器发来的请求发送消息时,将DICOM格式存储的图像转码为流媒体码,并以流码的形式传输到云代理服务器上。云代理服务器负责构建云节点系统,并采用分布式并行传输方式完成客户端与服务器端的通信任务。后端云数据库负责接收和存储来自于客户端的所有流媒体图像。

系统的工作流程如下:前端工作站在向云代理服务器发送请求的同时,也向客户端分发条消息;云代理服务器接到请求后,向前端工作站发回条确认消息,同时它又向云服务器分发条请求和套接字消息;后端云服务器接到请求后,回复条确认消息给云代理服务器,云代理服务器接到确认消息后,发送指令给客户端通知客户端发送图像文件。我们的系统是根据开源(Citrix XenServer)混合云设计的,它支持所有医学图像格式和视频格式的存储,并以流媒体形式在网络上传输。在此系统中我们构建了30个云节点,每个云节点彼此都以1Gbps的传输速率与ADSL网进行通信。

3.2 分布式并行文件传输设计

我们系统的软件设计采用的是基于Erlang分布式并行传输文件模式,Erlang有两种分布模式:一种是分布式Erlang,另一种是基于Socket的分布。我们的软件设计采用的是基于套接字的分布式模式。

服务器端程序代码如下:

```

1 %% ---
2 %% Erl3D Server - end
3 %% Programed by Wang Feng
4 %% ---
5 -modul(erl3d_server).
6 -compile(export_all).
7 -import(lists,[reverse/1]).
8 %% --- Parallelize
9 start_parallel_server() ->
10 {ok, Listen} = gen_tcp:listen(81, [binary, {packet, 4},
11                                     {reuseaddr, true},
12                                     {active, true}]),
13 spawn(fun() -> par_connect(, Listen) end).
14
15 par_connect(, Listen) ->
16 {ok, Socket} = gen_tcp:accept(Listen).
17 spawn(fun() -> par_connect(, Listen) end).
18 loop(Socket).
19
20 loop(Socket) ->
21 receive
22 {tcp, Socket, Bin} ->
23 [Filename | Value] = binary_to_term(Bin).
24 File_full_name = "d:/ErlangTest/" ++ Filename.
25 file:write_file(File_full_name, Value).

```

```

26. gen_tcp:send(Socket,term_to_binary(Filename)).
27. gen_tcp:close(Socket)
28. %%io:format("Server socket closed-n")
29. end.
30.

```

在程序中可以同时接受多个连接请求,来完成并行传输文件的任务,这比用单个服务器传输文件的通信速率提高了很多倍。同时图像文件以流媒体形式采用 TCP 连接协议进行传输也可以确保数据准确无误地发送到云数据库中。

3.3 云数据库虚拟化存储平台设计

后端云数据库存储平台由 Xen Server 服务器^[8]和存储设备等硬件资源构成,把物理服务器通过 Xen Server 的配置和整合就形成了可伸缩的虚拟化资源环境。Xen Server 的体系架构如图 3 所示。

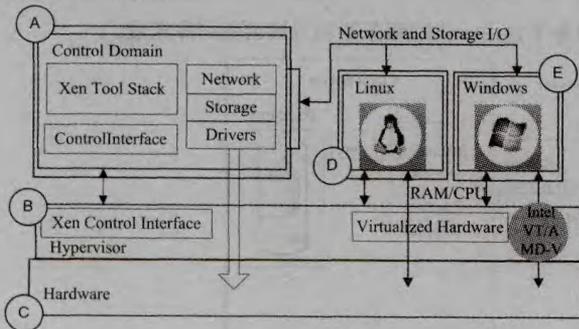


图 3 XenServer 体系架构

分别对图中 XenServer 体系架构中的不同组件作如下说明:

a) Control Domain (或称为 Domain 0) 是一个 Linux 虚拟机,对硬件而言,具有比来宾操作系统更高的优先级。Control Domain 管理所有来宾 VM 的网络和存储 I/O,而且它由于使用的是 Linux 设备驱动程序,因此能广泛支持各种物理设备。

b) Xen 虚拟机管理程序 (Hypervisor) 是运行于硬件上的一个软件薄层。Xen 提供一个允许每台服务器运行一台或多台“虚拟服务器”的抽象层,从而有效地将 OS 及其应用程序与底层硬件分离开来。

c) 硬件层包含物理服务器组件 (包括内存、CPU 和磁盘驱动器)。

d) Linux 虚拟机,其包括半虚拟化内核和驱动程序。通过 Control Domain 访问存储和网络资源,通过硬件上的 Xen 访问 CPU 和内存。

e) Windows 虚拟机,其使用半虚拟化驱动程序通过 Control Domain 访问存储和网络资源。Xen 经过设计可以充分利用 Intel VT 和 AMD-V 处理器虚拟化功能。硬件虚拟化可实现 Windows 内核的高性能虚拟化,而无需使用传统的仿真技术。

3.4 XenServer 管理设计

XenServer 的整体管理构架如图 4 所示。

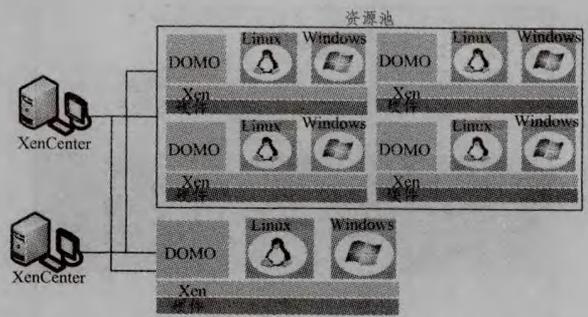


图 4 XenServer 的整体管理构架

我们借助资源池将多台虚拟化服务器作为单个实体进行管理。只需在一个位置执行认证,而不需要登录到所有的 XenServer。所有的服务器共享通用的网络和存储框架,从而方便使用自动 VM 放置和 XenMotion 等功能。

资源池采用主/从服务器管理模型,并将所有池配置数据复制到所有从属服务器上。这种配置可确保在主服务器发生故障时不会引起任何致命的故障。

XenCenter Client 作为图形化的管理控制台来集中管理 XenServer 服务器、虚拟机和资源。XenCenter 可以连接到多个服务器和资源池,且多个 XenCenter Client 可连接到同一个服务器或资源池。

XenServer 资源池中的 MetaData 会在各个服务器之间同步,不需要独立地控制管理中心,这样的架构防止了单点失败。缺省情况下 XenCenter 与资源池中的主控服务器连接,XenCenter 可以安装在常见 Windows 客户端操作系统中,通过物理服务器的管理 IP 地址与 XenServer 通讯。

4 实验与分析

4.1 实验环境及实验参数

在实验中,使用的实验环境如图 5 所示。在 HP785-G5 服务器上利用 Citrix XenServer 搭建了云计算平台,服务器上安装了 CentOS 5.3 企业版操作系统软件,使用了 30 个 VMS (虚拟机) 来构建 30 个云节点,云节点硬件配置是:CPU 是 1 Core; Intel X3330 2.66 GHz; 1 GByte Memory; Network 采用 1000 Base-T Gigabit Ethernet; Storage 是 3 Gbytes; Erlang OTP Ver. R14B03。我们用 Citrix Xen Center 软件来远程管理、操作和配置 VMS 的内部云服务器。

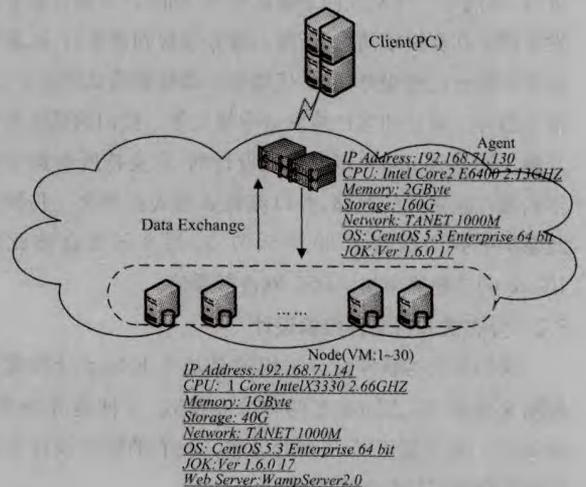


图 5 实验环境构架

分别在 Local to Local、TANet to TANet 和 ADSL to TANet 3 种网络环境下,对采用传统的 FTP(文件传输协议)传输的速率和分布式流媒体传输的速率进行比较。表 1 为 3 种类型的网络环境。

表 1 网络环境和带宽

	LOCAL	TANET	ADSL
bandwidth	1 Gbps	100 Mbps	2 Mbps

我们把从 CT 获得的 DICOM 图像文件进行传输。一共有 360 个图像(一个图像的每一度,从 0 到 360°)都需要重建 3D 图像文件。因此,在实验中,我们发送了 360 张 DICOM 图像文件,DICOM 图像文件信息如表 2 所列。

表 2 DICOM 图像文件信息

Medical Image file	
Format	DICOM 3.0
Type	CT image file
Resolution	640 * 640
Cross-sections	1
File Size	801mBytes

我们用流式传输和 FTP 传输进行实验,得到云的分布式存储和单服务器存储在传输时间上的不同,同时也用不同的节点数量来测试流媒体传输。每个节点上的 DICOM 的图像数量如表 3 所列。

表 3 节点上的 DICOM 的图像数量

Protocol	Nodes		
	1	5	30
FTP	360	*	*
Streaming	360	72	12

4.2 实验结果与分析

实验结果如表 4 所列,系统中有 1 个云节点时,相当于传统单一服务器的系统,在此系统中,可以看到局域网环境下的流媒体传输时间略小于 FTP 传输时间。在 TANET 网路和 ADSL 网路环境下传输流媒体的时间比 FTP 传输时间长,这说明流传输更稳定;系统中有 5 个云节点时,流的传输率是根据云节点数量的增加而增加的,它在 3 种类型网络环境下流的传输率分别增加 41%、51%和 13%;系统中有 30 个云节点时,流的传输率不继续增加,反而略有减少,这说明云节点数量增加到特定的值时,传输率也就达到最大值,云节点数量的特定值即是最佳值。当在系统中云节点的增加数量超过这个最佳值时,流的传输率不会随云节点数量的增加而继续提高。前客户端请求并行连接多个后端节点而造成的延迟,可能会增加传输时间。

表 4 实验结果

Network	Node	Streaming			FTP		
		1	5	30	1	5	30
Local		9.6±0.66	5.6±1.72	6±0.9	10.3±1.23	*	*
TANET		58.3±0.95	28.6±1.96	33.5±2.91	55.9±3	*	*
ADSL		633.8±9	533.1±2.23	598±10	580.1±4.78	*	*

注:时间单位为 s,数据格式是传输时间±时间标准偏差(做 30 个实验的数据)

比较局域网、TANET 和 ADSL 的实验结果可知:系统中

有 1 个云节点的情况,Local to TANET 网络环境下的流媒体传输时间从 9.6s 增加到 58.3s,FTP 传输时间从 10.3s 增加到 55.9s,即访问云服务器的时间增加近 6 倍;Local to ADSL 网络环境下的流媒体传输时间从 9.6s 增加到 633.8s,FTP 传输时间从 10.3s 增加到 580.1s,即访问云服务器的时间增加分别是 65 倍和 56 倍。

最后,可以看到系统中有 5 个云节点的流媒体传输时间与系统中有 30 个云节点的流媒体传输时间几乎一样。换句话说,系统中有 5 个云节点的流媒体传输足以支持 360 张 DICOM 图像文件的传输。

结束语 在本文中,我们研究了基于云的分布式存储和传统单一服务器存储,同时,在不同的网路环境下,也比较了流媒体文件传输的效率。实验结果表明,基于云医学图像流媒体传输平台的建立有效地解决了大量医学图像的存储和传输效率问题。采用虚拟化技术构建的云计算环境提供了快速和大规模部署虚拟机的管理员,从而使得云计算管理更容易。我们研究的结论是:

- (1)为了存储和传输大量医学图像文件,我们基于流媒体技术构建了一个可行的分布式存储架构。
- (2)在传统的单一服务器上传输图像文件时,用流媒体技术进行传输也可以减少图像文件传输时间。
- (3)实验结果表明,在增加云节点数量时,可以有效地减少数据传输和存储时间。对于文件传输率的问题,我们可以找到云节点数量增加的最佳值,当云节点数量超过了这个最佳值时,对提高文件传输率是无关紧要的。

参考文献

- [1] Cloud computing. Wikipedia[OL]. http://en.wikipedia.org/wiki/Cloud_computing
- [2] Bret M. In Clouds Shall We Trust[J]. IEEE Security & Privacy, 2009, 7(5): 3-3
- [3] 史佩昌,王怀民,蒋杰,等. 面向云计算的网络化平台研究与实现[J]. 计算机工程与科学, 2009(S1)
- [4] Duysburgh B, Lambrecht T, Turck F D, et al. An Active Networking Based Service for Media Transcoding in Multicast Sessions[J]. IEEE Trans. on System, Man, and Cybernetics-Part C: Applications and Reviews, 2004, 34(1): 19-34
- [5] Schulzrinne H, Casner S, Frederick R, et al. RTP: A transport protocol for real-time applications[Z]. 1996
- [6] PACS. Connecting for health[OL]. <http://www.connectingforhealth.nhs.uk/systemsandservices/pacs>
- [7] DICOM-Standard. NEMA[OL]. <http://medical.nema.org/standard.html>
- [8] Model of Distributed System. WECO Social Learning[OL]. <http://sls.weco.net/node/12902>
- [9] 夏英,杨选伦. 云环境中基于金字塔模型的影像数据存储方法[J]. 重庆邮电大学学报:自然科学版, 2012, 24(6): 669-674