

# 面向大规模汇聚接入网的分布式 VoD 服务的性能仿真

林 珉<sup>1</sup>, 王 新<sup>1</sup>, 郑常熠<sup>1</sup>, 汪斌强<sup>2</sup>, 薛向阳<sup>1</sup>

(1. 复旦大学计算机科学技术学院, 上海 200433; 2. 国家数字交换系统工程技术研究中心, 河南郑州 450002)

**摘 要:** 流媒体应用有着实时性、连续性、带宽消耗大等特点, 对传输网络提出了较高的要求. 用户接入网作为传输网络的重要组成部分, 对流媒体服务质量有重要影响. 新一代高性能宽带信息网的试运行, 初步验证了大规模汇聚接入网络技术在某些条件下支持流媒体的有效性. 然而, 由于某些限制, 在试运行中还不能实现在大规模用户并发访问流媒体服务器的条件下的验证, 而在仿真中却可以实现这种验证. 在本文中, 我们对于面向不同接入网的分布式 VoD 系统作了仿真比较. 结果表明, 面向大规模汇聚接入网的 VoD 系统在接入路由器无拥塞时, 在相同的硬件成本下, 在端到端时延及其抖动指标上有较大的提高.

**关键词:** 大规模汇聚; 接入网络; 视频点播; 分布式服务; 仿真

**中图分类号:** TP393      **文献标识码:** A      **文章编号:** 0372-2112 (2008) 10-2035-05

## Performance Simulation of Distributed VoD Service on Large Scale Converging Access Networks

LIN Min<sup>1</sup>, WANG Xin<sup>1</sup>, ZHENG Chang-yi<sup>1</sup>, WANG Bin-qiang<sup>2</sup>, XUE Xiang-yang<sup>1</sup>

(1. School of Computer Science, Fudan University, Shanghai 200433, China;

2. National Digital Switching System Research Center, Zhengzhou, Henan 450002, China)

**Abstract:** Streaming media services have characteristics of real time, continuity and large bandwidth consumption. These put a higher requirement on networks. An access network is an important part of internet and has significant effects on the service quality of streaming media. The test run on the China's High Performance Broadband Information Network has validated that the large scale converging access network can effectively support streaming media under some situations. However, due to some restrictions, the test run can not validated the efficiency under the situations when large scale users simultaneously access media servers. But the validations can be achieved in simulation. In this paper, we compare distributed VoD systems on different access networks by simulation. The results show that when the access router has no congestions, the VoD system on the converging access network has better performance in terms of end to end delay and jitter at the same hardware cost.

**Key words:** large scale converging; access network; video on demand (VoD); distributed service; simulation

## 1 引言

互联网上日益流行的视频点播 (VoD, Video on Demand)、电视直播 (Live TV) 等流媒体业务具有带宽需求高、传输实时和连续等特点, 要求传输网络在一定的时间内确保端到端的服务质量 (QoS, Quality of Service), 即足够大的带宽、足够小的延迟<sup>[1]</sup>. 由于缺少 QoS 保证, 目前互联网的突出表现就是不能有力支持流媒体业务. 尽管提出了很多措施<sup>[2]</sup>, 例如区分服务 (DiffServ, Differentiated Services)、集成服务 (IntServ, Integrated Services)、多协议标签交换 (MPLS, Multi Protocol Label Switching), 但是这些技术方案不能有效地解决问题, 而且还增大了路由器

的复杂性. 这些困难在传统因特网架构上 (IP 路由器之间的分布式互联, 复杂、无序、变化的拓扑) 是不能解决的. 因此, 仅仅依靠在路由器或用户终端上的处理, 无法从根本上保障 QoS, 网络的架构必须加以改变. 如何建立有效支持流媒体业务的网络架构已经成为当今互联网研究的热点之一.

### 1.1 相关工作

在国外, 美国是在开展下一代互联网研究和建设上最积极的国家. 有影响力的项目有 NGI、Internet2、Planet-Lab、MoonV6、100 × 100 和网络革新全球环境 (GENI, Global Environment for Network Innovations) 等. 其中, 100 × 100 项目的核心思想是<sup>[3]</sup>: 未来网络应该抛弃路由器到

路由器互联的分布式构架,消除传统因特网的缺陷,致力于构建全新网络,而不考虑与现有网络的兼容。GENI 计划也认为<sup>[4]</sup>,未来网络的发展应该是革命性的,而不是演化式的。斯坦福大学博士 Fernandez 则提出<sup>[5]</sup>:如果重新构建网络,需要做到在网络的边缘进行分组交换,在网络的核心进行电路交换,并把这两种交换集成起来。

在我国,重要的网络计划有下一代网络<sup>[6]</sup>(NGN, Next Generation Network)、下一代互联网示范工程<sup>[7]</sup>(CNGI, China Next Generation Internet)、高性能宽带信息网<sup>[8]</sup>(3TNet)等。其中,3TNet 是国家 863 计划对下一代网络研究的重大专项,其设计采用分层的结构,在核心网络采用基于电路交换的自动交换光网<sup>[9]</sup>(ASON, Automatic Switched Optical Network),在边缘接入网采用 IP 分组交换,通过简化接入网拓扑,提供端到端的 QoS 保证。

## 1.2 我们的贡献

虽然 3TNet 的试运行初步验证了大规模汇聚接入网在某些条件下支持流媒体的有效性,但是由于某些原因,它无法实现在大规模用户并发请求 VoD 服务条件下的验证,而仿真可以克服这个困难。在本文中,我们用 OPNET 仿真实现了这种验证。其主要特点是:模拟了 5 万用户同时访问 VoD 服务;根据路由器上有无阻塞分两种情况仿真;以端到端时延、时延抖动和丢包率作为评估指标,分析比较了三种不同拓扑接入网上的 VoD 性能。仿真验证了:汇聚接入网在大规模并发 VoD 服务的压力下,如果接入路由器提供足够的处理与存储能力,在用户接收端上的 VoD 收看效果可以得到改善。而且在同样的性能需求下,大规模汇聚接入网更节省硬件成本。

## 2 支持流媒体的大规模汇聚的架构

### 2.1 支持流媒体的网络架构

传输网络按照不同的功能从结构上可以分为核心骨干网与用户接入网两大部分。核心骨干网是指处于整个网络中心的用于把位于网络边缘上的各个部分网络连接起来的骨干部分。其作用就是为各种边缘网络之间的通信提供高速、可靠的传输信道。用户接入网部分是指处于网络边缘上的直接为各种类型的终端设备提供接入服务的网络。在边缘网络上一般采用分组交换以实现统计复用。同时,为了保证大规模并发实时流媒体的端到端 QoS,在接入网就需要提供足够的宽带和加速比,所以在接入网部分应该提供一种高度简化的架构。

### 2.2 大规模汇聚接入网的架构

图 1 是大规模汇聚接入网架构。其最大特点是接入网内用一台高性能路由器同时为大规模的家庭或企业用户提供 IP 接入服务<sup>[10]</sup>。这种拓扑实现了最大可能的

简化,即从传统的复杂拓扑(即用若干台不同规模的 IP 路由器和二层交换机等互联,形成复杂无序的拓扑)简化为树形拓扑,以大规模接入路由器为树根节点实现 IP 交换,同一个接入网中任意两个用户之间的交互都要经过接入路由器,且传输路径只有 2 跳(hop)。

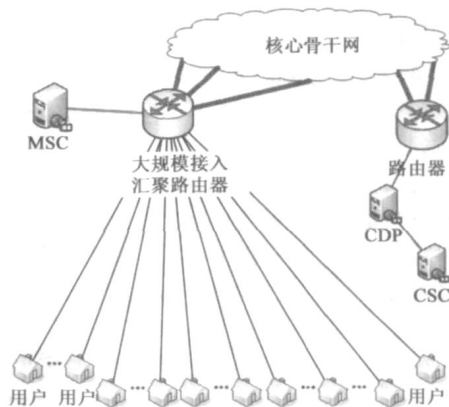


图1 基于大规模汇聚接入网的分布式VoD系统

## 3 面向大规模汇聚接入网络的 VoD 服务模型

### 3.1 组成部分

图 1 是 3TNet 专项为大规模汇聚接入网设计的一种分布式 VoD 系统。它主要由三个部分组成:(1)内容服务器集群(CSC, Content Server Cluster), CSC 具有海量的磁盘空间,用于存储由电视台、电影公司等内容提供商所提供的所有视频节目;(2)内容分发平台(CDP, Content Distribution Platform), CDP 是一组服务器集群,负责把影片分发到媒体服务器上。分发的对象通常是那些经过选择的近期最受欢迎的影片,并根据影片类别决定是否采用组播分发;(3)媒体服务器集群(MSC, Media Server Cluster), MSC 可以看作是接入网络上的本地 VoD 服务器,为接入汇聚路由器所连接的数以万计的用户提供实时 VoD 业务。

### 3.2 基本工作原理

在系统起始阶段, CSC 把热门影片传输给 CDP, CDP 利用骨干网提供的链路以组播的方式分发到各个 MSC 上。在绝大多数情况下, MSC 以代理服务器的方式工作,响应用户所有 VoD 请求。如果在 MSC 上已经存储所请求的影片,系统会根据调度策略播放相应的视频流。否则, MSC 会发起一次请求,从 CSC 上通过 CDP 将影片以突发的方式下载到 MSC 上。由于这个系统是基于 T 比特级传输设计的,所以从 CSC 下载一部高画质的影片到 MSC 上只要数秒时间。因此, VoD 的服务质量,特别是影片启动延迟,不会受到太大影响。

当骨干网负载较轻时, CSC 会把新影片和热门影片通过 CDP 分发到 MSC 上, MSC 上存储的其他影片则根据用户请求从 CSC 下载获得。MSC 的工作方式类似于

计算机系统中的高速缓冲(Cache),当 MSC 存储空间不足并且有新影片需要存储的时候将采用最近最少使用(LRU, Least Recently Used) 替换策略来决定替换的影片.

### 3.3 主要优点

大规模汇聚网络架构下的 VoD 服务模型相比其他网络架构具有以下这些优点: (1) 这种 VoD 系统具有简化的拓扑: 从 MSC 到用户终端只需要经过 2 跳, 既减少了传输代价, 又能降低接入延迟和传输抖动; (2) 由于人类行为具有相似性, 大量用户群中不同用户同时点播同一部影片的概率会增大, 所以 MSC 可以在接入网上利用组播对视频流聚链; (3) 管理扩展性强: 虽然 MSC 设计成本很高, 但是在大规模使用的情况下每户用户的平均成本是合理的. 当网络规模增大、用户增多时, 可以通过增加新的接入汇聚路由器模块来扩充网络, 新增用户的 VoD 请求将由新的 MSC 来负责服务, 通过这种可扩展的架构来保证端到端的 QoS.

## 4 性能评估

### 4.1 仿真环境

我们用 OPNET 对基于不同接入网的 VoD 系统作了仿真比较. 主要的环境设置如下:

#### (1) 接入网拓扑

我们比较三种接入网, 下面分别记为 Topo1、Topo2、Topo3. Topo1 是大规模汇聚接入网(图 1); Topo2 是由一个路由器和若干交换机组成的树形接入网(图 2a); Topo3 是由若干路由器组成的网状接入网(图 2b). 为了简明起见, 在图 2 中省略了 CSC 和 CDP. 选择后两种接入网作为比较对象, 是因为它们是传统接入网的典型代表. 为便于比较, 这三种接入网上路由器、交换机、服务器、链路速率都采用相同的配置.

表 1 三种接入网的架构比较

类型	汇聚接入	拓扑复杂度	拓扑特点	分组复用度	用户间跳数	路由器个数	交换机个数
Topo1	是	低	2 层树形	高	2 跳	1 个	0 个
Topo2	否	中	多层树形	中	不定	1 个	多个
Topo3	否	高	连通网状	低	不定	多个	多个

#### (2) VoD 业务模型

仿真中只关注在接入网内部的 VoD 性能, 而忽略骨干网上的传输. 这是因为用户的请求绝大多数都是从本地接入网的 MSC 上得到服务的, MSC 作为 CSC 的本地代理可以提供很高的点播命中率.

我们设定每个接入网上包含 5 万用户. MSC 以类似于传统集中式服务器系统的方式工作, 对本地接入网上的用户点播请求进行响应. 我们不考虑 MSC 拒绝请求的情形. 用户请求相互独立, 服从泊松分布(poisson), 以参数  $\lambda$  代表用户请求到达率<sup>[11]</sup>.

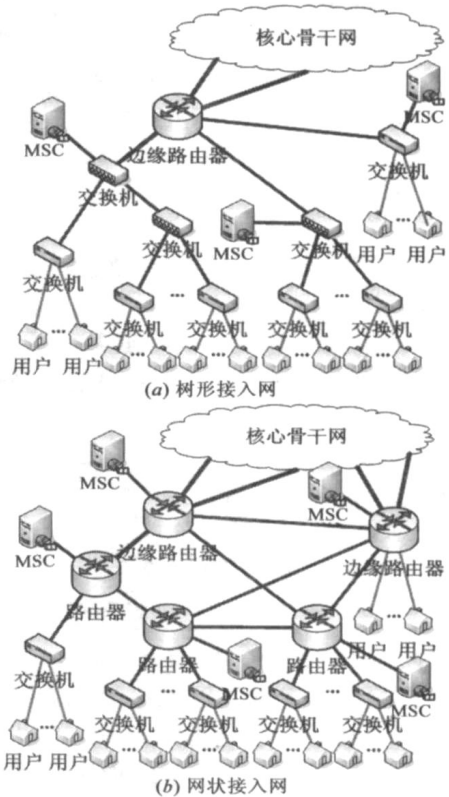


图 2 面向不同接入网的 VoD 系统

我们还设定 MSC 上存储 100 套 VoD 影片. 影片点播的热门程度服从参数为 0.271 的 Zipf 分布. 用户在线观看时间通过 VoD 影片的长度来模拟<sup>[12]</sup>. 影片的时间长度(分钟)服从对数正态分布: 平均值  $\mu = 60$ , 方差  $\sigma = 17$ . VoD 影片以可变比特率(VBR, Variable Bit Rate)方式传送, 传输层采用 UDP 协议(User Datagram Protocol).

### 4.2 性能评估指标

VoD 服务流是以实时视频流形式在网络中传输的. 不同于传统的文件传输, VoD 从传输开始起, 就必须以稳定的速率发送到用户终端上, 以确保视频的平滑播放, 尽量不能停顿和间断. 处理器竞争或带宽瓶颈等都会引起传输的延迟, 造成视频流的阻塞.

由于用户端感知的主观满意程度是 VoD 业务对网络的基本要求, 所以仿真中我们将着重关注 VoD 的端到端时延、时延抖动和丢包率等指标.

(1) 丢包率是指系统总丢包数与总发包数之比. 由于网络拥塞, 路由器上丢弃部分数据包, 使用户接收不到完整的信息, 在解码过程中出错, 用户观看时会出现马赛克.

(2) 端到端时延是指视频包从 MSC 发送到用户所需的时间, 这个时间应当尽量短.

(3) 时延抖动是指端到端时延变化的剧烈程度, 抖动过大将影响收看的实时性.

### 4.3 仿真结果及分析

根据路由器上有无拥塞, 我们分两种情况仿真. 1、无拥塞时, 我们把所有网络资源设为足够多; 2、有拥塞时, 我们把路由器上的资源设为紧缺, 但其它资源仍然设为足够多.

#### (1) 无拥塞的情况

我们仿真了三种接入网的平均端到端时延随用户请求到达率 $\lambda$ 的变化情况. 路由器转发速率为 225Mbps. 图3表明, 无拥塞时平均端到端时延基本上不随 $\lambda$ 变化. 其中 Topo1 上的平均端到端时延最小. 因为路由是网络时延的主要因素, Topo1 上的端到端传输跳数最小, 即只有 2 跳, 因此时延最小.

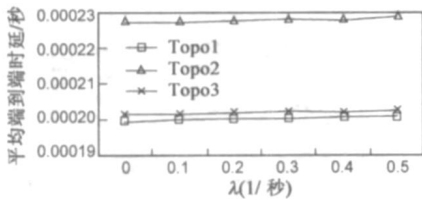


图3 无拥塞时, 平均端到端时延随 $\lambda$ 的变化

我们还仿真了三种接入网在 $\lambda$ 分别为 0.05 和 0.5 时的端到端时延随时间的变化情况. 图4表明, 在端到端时延的抖动上, Topo1、Topo3 较小, 而 Topo2 较大.

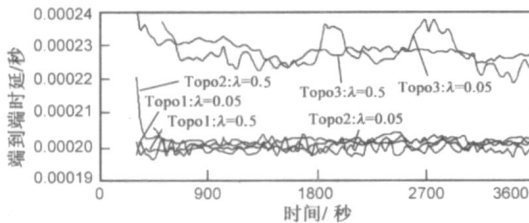


图4 无拥塞时, 端到端时延随时间的变化

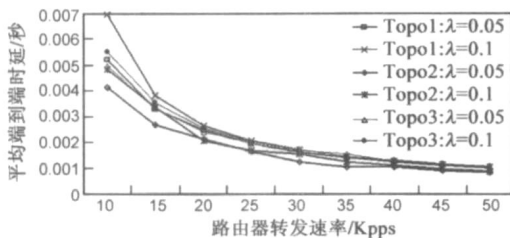


图5 无拥塞时, 平均端到端时延随路由器转发速率的变化

我们仿真了三种接入网在 $\lambda$ 分别为 0.05 和 0.1 时的平均端到端时延随路由器转发速率的变化情况. 图5表明, 各种拓扑的平均端到端时延随着路由器转发速率增大而减小, 且它们之间的差距也在减小. 另一方面, 由于接入网的硬件成本主要由路由器和交换机组成, 且各个路由器或交换机的性能参数设为都相同. 所以, 在取得相同服务性能时 Topo1 所需的硬件成本最小.

(2) 有拥塞的情况我们仿真了三种接入网的丢包率和平均端到端时延随 $\lambda$ 的变化情况. 图6表明, 对于这三种接入网中的任一种, 有拥塞时丢包率都随 $\lambda$ 增加而

增加. 其中 Topo1 上的丢包率最大. 图3和图7表明, 任一种接入网上的平均端到端时延, 有拥塞时比无拥塞时大得多.

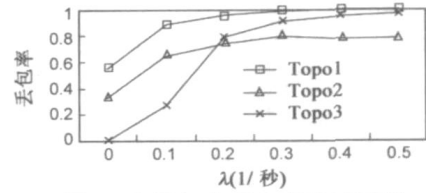


图6 有拥塞时, 丢包率随 $\lambda$ 的变化

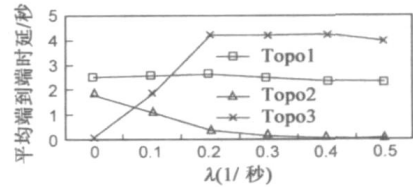


图7 有拥塞时, 平均端到端时延随 $\lambda$ 的变化

总之, 基于大规模汇聚接入网的 VoD 系统, 在无拥塞时, 在相同硬件成本下其端到端时延及其抖动的指标有较大提高. 但是, 在有拥塞时, 这种系统在丢包率和端到端时延上显著变差. 这是因为这种系统的接入路由器要汇聚本地接入网的所有流量, 所以要承受更大的负载, 从而更易形成单结点瓶颈.

为了避免拥塞, 可以采取以下几种措施. 一是采用高性能的接入路由器, 以提供足够强大的处理与缓存能力, 同时提供很高的可扩展性, 以满足业务不断发展的需求. 二是在接入路由器上采用组播技术, 以进一步缓解路由器上的流量压力.

### 5 结论

目前, 通信网络的发展趋势是不同业务的融合. 流媒体业务具有并发、组播和高宽带特性, 是融合业务中的关键和难点, 因此通过设计合理的网络架构和服务模型, 研发高效的设备和业务软件, 可以从根本上满足流媒体的需求, 并扩展未来的新业务, 这具有十分重要的意义. 在本文中, 我们用 OPNET 实现了对面向不同架构的接入网的分布式 VoD 系统在大规模用户条件下的仿真比较. 仿真表明这种 VoD 系统在无阻塞时, 在降低接入网硬件成本的前提下, 在端到端时延及其抖动等指标上有所改善.

#### 参考文献:

- [1] 纪其进, 董育宁. IP 网络性能特征模型分析 [J]. 通信学报, 2004, (03): 153- 162.
- [2] 袁捷, 张同须. IP 网络及其 QoS 技术发展研究 [J]. 电信科学, 2005, (05): 27- 31.
- [3] Hui Zhang, et al. Collaborative research: ITR/ ANIR: 100Mb/sec for 100 million households [ DB/OL ]. <http://www.cnki.net>

- 100x100network. org/papers/100x100proposal. pdf, March, 2003.
- [ 4 ] Tom Anderson, et al. GENI: Global Environment for Network Innovations, Conceptual Design Project Execution Plan[ DB/OL]. <http://www.geni.net/GENI10JAN06.pdf>, January 10, 2006.
- [ 5 ] Pablo Molinero Fernandez. Circuit Switching in the Internet [D]. USA: Stanford university, 2003. 46- 47.
- [ 6 ] 赵慧玲, 徐向辉. NGN 的研究进展[ J]. 电信科学, 2004, (01): 34- 39.
- [ 7 ] 蒋林涛. CNGI——中国下一代互联网示范项目[ J]. 数据通信, 2005, (01): 21- 22.
- [ 8 ] 邬江兴. 中国高性能宽带信息网(3TN et) 综述[ J]. 通讯世界, 2002, (12): 14- 16.
- [ 9 ] 雷震洲. 核心网中的光分组交换[ J]. 电信工程技术与标准化, 2003, (01): 7- 11.
- [ 10 ] 汪斌强, 邬江兴. 基于 IPv6 的大规模接入汇聚路由器的设想和实现[ J]. 电信科学, 2006, (01): 10- 14.
- [ 11 ] Hongliang Yu, Dongdong Zheng, Ben Y Zhao, Weimin Zheng. Understanding User Behavior in Large Scale Video Demand Systems[ A]. In Proceedings of the 2006 ACM 1st Eurosys Conference[ C]. European: EuroSys, April 2006. 333- 344.
- [ 12 ] 罗建光, 赵黎, 杨士强. 基于用户行为分析的应用层组播树生成算法[ J]. 计算机研究与发展, 2006, (09): 1557- 1563.

#### 作者简介:

林 珉 男, 1976 年 5 月生于江西省新余市, 复旦大学计算机科学与工程系博士生, 主要研究方向为支持流媒体的网络架构等.

王 新 男, 1973 年 8 月生于陕西省三原县, 复旦大学计算机科学与工程系副教授, 主要研究方向为多媒体网络. (本文通信作者)

E-mail: xinw@fudan.edu.cn

郑常熠 男, 1983 年 7 月生于福建省宁德市, 复旦大学计算机科学与工程系硕士研究生, 主要研究方向为支持流媒体的网络架构等.

汪斌强 男, 1963 年 2 月生于安徽省安庆市, 河南省郑州市国家数字交换系统工程技术研究中心教授, 博士生导师, 主要研究方向为宽带信息网络与高速路由器核心技术等.

薛向阳 男, 1968 年 6 月生于江苏省建湖县, 复旦大学计算机科学与工程系教授, 博士生导师, 主要研究方向为多媒体信息检索和处理等.