

混合云可扩展视频编码的视频直播机制研究^①

周传华, 江超, 赵伟

(安徽工业大学 管理科学与工程学院, 马鞍山 243032)

摘要: 随着网络技术的不断发展, 使得网络视频的应用研究成为新的研究热点. 针对用户请求异构性和服务器资源分配的问题, 本文对传统的视频直播技术架构及存储和调度机制进行研究, 提出了一种基于混合云存储的可扩展视频编码视频直播存储及调度机制. 仿真数据表明, 该机制能明显降低视频直播系统中的网络延迟.

关键词: 混合云; 视频直播; 服务器选择; 可扩展视频编码

引用格式: 周传华, 江超, 赵伟. 混合云可扩展视频编码的视频直播机制研究. 计算机系统应用, 2017, 26(7): 258-262. <http://www.c-s-a.org.cn/1003-3254/5832.html>

Live Video Mechanism of SVC Based on Hybrid Cloud

ZHOU Chuan-Hua, JIANG Chao, ZHAO Wei

(School of Management Science and Engineering, Anhui University of Technology, Maanshan 243032, China)

Abstract: The application of network video has become a new research hotspot with the development of network technology. After analyzing the traditional technical architecture of live video broadcast and the mechanism of storage and dispatch, a storage and distribution mechanism of scalable video coding is proposed based on hybrid cloud in this paper in view of the problem of both the heterogeneity of user request and resource allocation. Finally, conclusions are confirmed through the simulation data that the mechanism can significantly reduce the network delay in the video streaming system.

Key words: hybrid cloud; live video; server selection; scalable video coding

1 引言

近年来, 随着多媒体技术和网络技术的飞速发展, 基于网络的视频直播技术被广泛的应用. 如何更好的利用Internet为人们提供高质量的视频服务, 成为当下研究的热点. 在大规模服务器集群系统中, 对视频数据的科学存储方式以及对服务请求的合理调度是提高系统性能的关键, 传统的视频直播系统主要研究如何均衡服务器的负载问题.

刘志强^[1]等学者对云环境中的视频直播服务器选择问题进行了研究, 将服务器选择问题归结为一个优化问题, 提出了一种多阶段调度方法. 刘景^[2]提出一种基于云计算的视频直播设计方法, 该方法采用分层的

建设思路动态调配资源建立分发服务器, 有效地节约了校园网主干线路的大量带宽和物理计算资源. 孙名松^[3]等学者为了提高校园网视频直播系统使用性能, 提出了一种基于云计算的视频直播原型系统, 该系统解决了单点故障隐患, 重点改善了负载不均衡问题, 提高了服务器利用率. 张欣^[4]等学者设计了一种用于视频直播的实时流媒体负载分配方案. 于雷^[5]在分析校园网网络直播系统的需求基础上, 进行了网络直播系统方案设计. 刘侃^[6]针对共享存储结构下的服务器机群, 提出一种就与视频节目点播集中度的调度策略, 该策略能有效提高视频服务器集群的运行性能. 罗建光^[7]等学者介绍了一种基于P2P网络的大规模视频直播系统Gridmedia, 该系统采用Gossip协议构建无结构的应用层覆盖网络,

^① 基金项目: 安徽省科技厅重大软科学项目(1502052006)

收稿时间: 2016-10-13; 收到修改稿时间: 2016-11-21

每个节点可以独立地选择自己的伙伴节点. 在覆盖网络上, 每个节点通过一种推拉结合的流传输策略从邻居节点获取数据, 有效降低了直播系统的控制开销. 秦丰林^[8]等学者深入介绍和分析了覆盖网络构建、数据块调度策略等P2P流媒体核心技术, 并概要讨论视频编码、网络编码、覆盖网络拓扑优化、安全与激励机制等P2P流媒体相关增强技术, 并对将来研究方向作了展望. 彭宏^[9]等学者, 介绍了一个利用3G/4G手机进行视频拍摄和传输的移动直播系统. 分析了当今移动直播系统的发展趋势, 介绍了如何设计并实现一个基于流媒体技术的移动视频直播系统. 文献^[10,11]介绍了P2P视频直播系统的相关知识, 对比了当前主流P2P视频直播系统的主要结构, 针对现有大规模直播系统的不足, 提出了一些改进. Hajjat M^[12]等学者提出了一种动态服务器选择策略, 能够更加合理的利用服务器资源, 减小延迟. Grandison T^[13]等学者基于云资源的虚拟资源池提出一种动态分配负载策略, 以达到高的资源利用率.

以上研究均没有关注到视频数据的编码方式和编码后的视频数据如何存储在服务器中以及服务器如何进行视频数据的分发等方面的研究. 在视频直播环境中, 科学的编码方式以及合理的存储及分发机制对于端到端的延迟有着较大的影响, 是视频直播系统中关键的因素之一, 对于用户端的QoS有着重要的影响. 因此, 本文提出一种基于混合云存储的网络环境, 采用可扩展视频编码的方式对直播视频数据进行存储及分发的机制.

2 模型构建

2.1 相关技术

云存储^[14](cloud storage)是在云计算的概念上延伸和发展出来的一个新概念, 它是通过集群应用、网络技术或分布式文件系统等功能, 将网络中大量各种不同类型的存储设备通过应用软件集合起来协同工作, 共同对外提供数据存储和业务访问功能的一个系统. 混合云是公有云和私有云两种服务方式的结合, 是目标架构中公有云和私有云的结合. 混合云充分利用了公有云和私有云的优势, 既发挥了公有云价格低廉、扩展性强的特点, 又有了私有云服务性能好等优势, 这使得混合云的应用变得非常广泛. 本文正是基于混合云的优势展开了实时视频存储及调度策略的研究.

在视频直播应用中, 合理的编码方式对于提高系

统的存储效率与对用户请求的快速响应有着非常重要的意义. 由于用户终端和网络异构性使单纯追求高压缩率的编码方式很难满足用户需求, 解决这一问题的方法是采用可扩展视频编码^[15](scalable video coding, SVC). 可扩展视频编码是H.264/AVC基础上扩展出来的一种技术, 其目标具有H.264相似的编码效率、相似的解码复杂度和相似的重建质量, 它拥有H.264/AVC的全部特点, 但增加了可扩展功能. 可扩展视频编码技术包含多个方面的可扩展性, 其中包括重要的3个方面: 时间可扩展、空间可扩展和质量可扩展.

可扩展视频编码技术可以将视频编码为一个基础层(base layer, BL)和若干个增强层(enhancement layer, EL). 基础层的数据可以在客户端进行解码而获得基本的视频质量, 而增强层数据可以从空间、时间和质量三个维度对基础层视频质量进行扩展, 解码的增强层数据越多获得的视频质量也就越高. 基于可扩展视频编码的视频分层模型如图1所示.

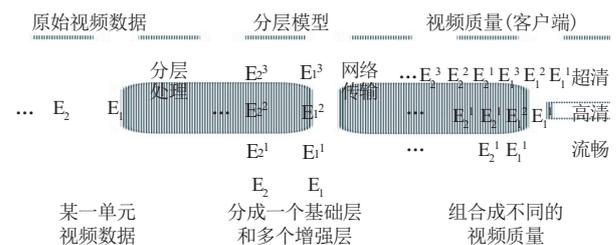


图1 可扩展视频编码在视频中的应用

在该分层模型中, 将原始视频数据的某一单元进行分层编码处理, 从而获得多层数据, 分为基础层和多个增强层. 用户端在通过网络获取这些数据后, 重新进行解码组合, 分为以下几种情形: (1)只收到基础层 E^1 的数据, 视频质量只能达到基本保证(表现为流畅); (2)如果在视频延迟要求范围内收到基础层 E^1 和增强层 E^2 的数据, 用户端可以获得更好的视频质量(高清); (3)如果在视频延迟要求范围内收到基础层 E^1 、增强层 E^2 和增强层 E^3 的数据, 用户端可以获得超清的视频质量. 如果用户端收到基础层 E^1 、增强层 E^3 的数据, 用户端也只能获得流畅的视频质量. 所以用户想要获得更高的视频播放质量, 必须获得该视频类型所包含的最高层之前的全部视频层数据.

2.2 系统模型

2.2.1 系统组成

视频直播系统由视频采集端、编码器、视频服务

器、混合云服务器组成. 视频源端的摄像头以固定的帧率进行视频采集, 编码器对视频数据进行可扩展视频编码, 并将编码后的数据进行压缩和RTP封装以提高视频传输的实时性. 资源调度中心将编码处理后的视频数据发送到视频服务器, 视频服务器负责接收或者拉取来自编码器的数据并将这些数据转发和存储到各地区的公有云服务器和私有云服务器. 系统模块如图2所示.

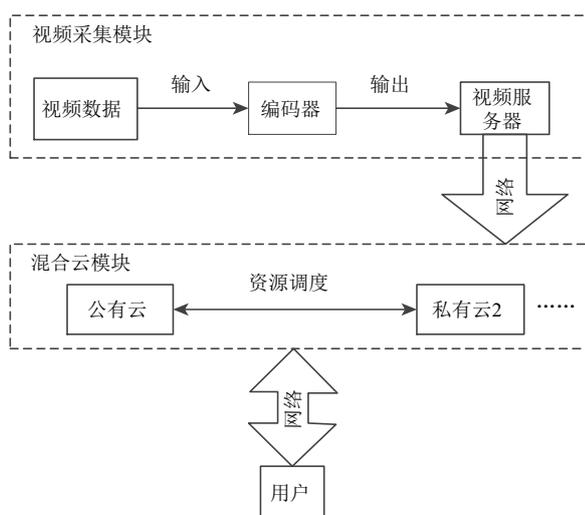


图2 系统模块图

2.2.2 服务器视频资源分发机制

科学的视频资源分发机制将有利于提高服务器的资源利用率, 提高视频直播的实时性. 资源调度中心在选择使用公有云存储资源时, 要支付相应的费用, 使用的公有云存储资源越多, 需要支付的费用也就越多, 从而增加了运营成本. 而且公有云存储资源的计算能力也不及私有云存储资源, 所以本机制将最小化公有云存储资源的使用. 本机制使用的私有云存储资源(edge cloud)^[16]的计算能力相对优于公有云存储资源, 且其使用成本远小于公有云存储资源, 所以本机制将最大化私有云存储资源的使用. 但私有云存储资源具有稳定性差的缺点, 例如提供服务的私有云存储资源突然停机等, 使得用户端接收的视频质量得不到保障. 所以本机制针对公有云存储资源稳定性高、成本高以及私有云存储资源稳定性差、成本小的特点, 将使用公有云存储资源存储重要性最高的基础层视频数据, 分别使用数量较多的私有云存储资源存储重要性较小的增强层1和增强层2视频数据, 并对私有云存储资源使用备份机制, 从而增加一定的成本来保障更好的服务质量.

在本机制中, 视频服务器将编码后的基础层视频数据通过覆盖网络(一般为专用网络, 也可以为公共网络)发送到各地的公有云服务器, 将增强层1和增强层2的视频数据发送到各地区的私有云服务器. 由公有云服务器为发送视频请求的用户提供基础层视频数据流, 以保障用户能得到最基本视频质量的直播视频, 由私有云服务器为发送视频请求的用户提供增强层1或者增强层2的视频数据流. 视频请求为流畅的用户将得到基础层的视频流, 此用户请求将被重定向到该区域中距离用户最近的一台公有云服务器; 视频请求为高清的用户将得到基础层和增强层1的视频流, 此用户请求将被重定向到该区域中距离用户最近的一台公有云服务器和一台提供增强层1视频流的私有云服务器中; 同理, 视频请求为超清的用户将得到基础层、增强层1和增强层2的视频流, 此用户请求将被重定向到该区域中距离用户最近的一台公有云服务器、一台提供增强层1视频流的私有云服务器和一台提供增强层2视频流的私有云服务器中.

2.2.3 服务器视频资源分发动态调整机制

资源调度中心对接入系统中的用户进行实时监控, 计算出该区域中用户请求类型为流畅、高清和超清的用户比例. 根据用户比例调整接入该区域中提供增强层1的私有云服务器和提供增强层2的私有云服务器的数量比例. 正在提供服务的服务器一旦确定提供什么样的视频资源, 在系统正常工作时就不会改变该服务器来调整服务器比例, 而是通过调整新加入服务器提供相应的视频数据来调整系统中接入的服务器的数量比例关系. 这样做的目的是为了保障系统的稳定性, 不至于因为改变正在工作的服务器而使系统中正在接受服务的用户在接收视频数据时受到影响而降低视频质量. 当接入系统中的服务器不提供服务时, 资源调度中心将释放该闲置的服务器, 防止资源的浪费.

本文机制的目的为在满足用户请求和服务器负载均衡的条件下最小化直播系统中的网络传输延迟. 如图3所示.

2.2.4 数学建模

资源调度中心在将用户请求重定向到相应的服务器前, 首先要判断该服务器是否满足可以为当前用户提供服务的条件, 如果满足条件, 则将该用户请求重定向到该服务器; 如果不满足条件, 则将该用户请求重定向到该区域中相同类型的服务器中. 判断接入系统中

的服务器是否满足可以为用户提供服务的条件时需要考虑以下几个方面因素^[1]: 网络接口带宽、内存和CPU处理资源。

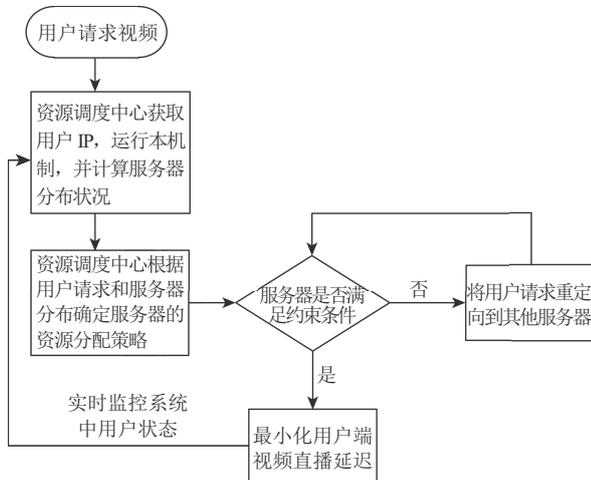


图3 基于混合云存储SVC视频直播机制流程

网络接口带宽只有满足一定条件时, 服务器才能正常提供服务. 设 B_{total} 为网络接口总带宽, B_{used} 为已经使用的带宽容量, α 为带宽因子, 该值与网络环境和操作系统有关, 则只有满足 $B_{total} \cdot \alpha > B_{used}$ 时, 服务器才能正常为用户提供发送视频流的服务. 假设公有云服务器的网络接口带宽为 B_p , 私有云服务器的网络接口带宽为 B_{pr} , 则应有如下关系:

$$C_1 \times n_1 \leq B_p \times \alpha \quad (1)$$

$$C_i \times n_i \leq B_{pr} \times \alpha, i = 2, 3 \quad (2)$$

式中, C_1 为服务器为用户端发送基础层视频流所占用的带宽, C_2 和 C_3 分别为服务器为用户端发送增强层1和增强层2视频流所占用的带宽. n_1, n_2 和 n_3 分别为接入到公有云服务器、提供增强层1的私有云服务器和提供增强层2的私有云服务器中的用户请求数量.

接入服务器中的用户的数量越多, 服务器被消耗的内存容量也随之增加. 假设服务器提供不同的视频数据流所消耗的内存量是不同的, 则对于提供单一视频数据的服务器的内存消耗量与接入该服务器的用户量成线性关系. 假设 R_p 为公有云服务器的内存容量, R_{pr} 为私有云服务器的内存容量, β 为内存比例因子, 该值的取定与操作系统、媒体播放平台有关, 则应有如下关系:

$$R_1 \times n_1 \leq R_p \times \beta \quad (3)$$

$$R_j \times n_j \leq R_{pr} \times \beta, j = 2, 3 \quad (4)$$

式中, R_1, R_2 和 R_3 分别为服务器发送基础层数据流、增强层1数据流和增强层2数据流所消耗的内存.

另一方面, 服务器的CPU资源也要满足一定的约束. 假设每个接入系统中的用户会占用 τ_i 的CPU资源, 则CPU的使用率可以表示为 $\theta_i^{CPU} = \lambda_i + x_i \cdot \tau_i$, 其中 λ_i 为CPU背景负载, x_i 是当前接入服务器中的用户数量. 假设服务器的响应时间 g_i 为CPU使用率的线性函数, 则有 $g_i = a_i \cdot \theta_i^{CPU} + b_i$, 其中 a_i 和 b_i 为常数. 服务器的响应时间对用户体验QoS有着重要的影响.

在满足(1)~(4)式的情况下, 资源调度中心将最小化用户端对于直播视频的播放延迟. 则本文提出的基于混合云存储的视频直播的存储和调度机制的效益函数为:

$$\min(a \sum_{i=1}^{N_1} d_i + b \sum_{i=1}^{N_2} g_i) \quad (5)$$

式中, N_1 为接入系统中的用户端数量, N_2 为系统中提供服务的服务器数量, d_i 为第 i 个用户端所接入的所有服务器中的最大网络传输延迟. a 和 b 为权重因子, 可以根据需要对 a 和 b 设置大小关系: 当 $a > b$ 时, 重点关注直播系统中网络环境的差异性, 此时该机制适用于网络性能差别较大的直播系统; 当 $a < b$ 时, 重点关注直播系统中服务器对用户请求的响应延迟, 此时该机制适用于混合云中服务器性能差别较大直播系统.

3 性能分析

本文将从理论上对提出的机制进行分析, 通过matlab R2015b仿真软件对使用该机制所产生的延迟进行仿真. 用户端获取直播视频过程中的延迟主要包括服务器对视频流数据的处理延迟、视频流数据在网络中的传输延迟和用户端对接收到的视频流进行处理的延迟. 本文主要研究服务器对视频流数据的处理延迟和视频流数据在网络中的传输延迟, 用户端对接受的视频流的处理延迟可作为常量.

直播视频资源在服务器端的处理可用排队论的知识进行表达, 其中公有云服务器处理能力可用参数为 u_1 的指数分布进行表达; 同理私有云服务器处理能力可用参数为 u_2 的指数分布进行表达. u_1 和 u_2 满足以下关系: $u_1 > u_2$. 系统中的用户请求量可用参数为 λ 泊松分布进行表达. 数据在网络中的传输延迟服从正态分布, 假设视频流数据在公有云服务器到用户端的网络传输延

迟可用参数为 u_3 的正态分布进行表达,同理,视频流数据在私有云服务器到用户的网络传输延迟可用参数为 u_4 的正态分布进行表达.仿真结果如图4所示.

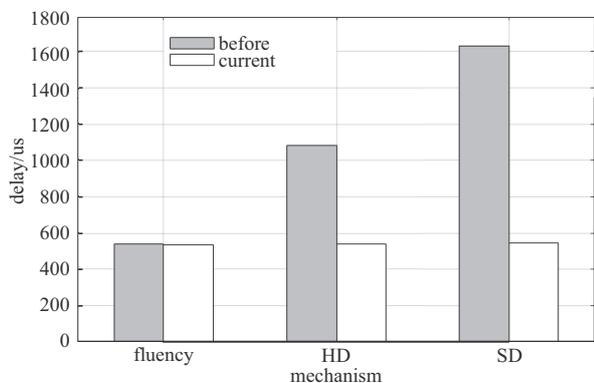


图4 使用本文机制前后所产生的延迟对比

图4的结果如表1和表2所示.

表1 使用本文机制前所产生的延迟

| 用户请求类型 | 流畅 | 高清 | 超清 |
|--------|--------|---------|---------|
| 延迟(us) | 537.63 | 1081.08 | 1630.43 |

表2 使用本文机制后所产生的延迟

| 用户请求类型 | 流畅 | 高清 | 超清 |
|--------|--------|--------|--------|
| 延迟(us) | 534.80 | 540.54 | 543.48 |

表1展示了使用本机制前直播系统产生的延迟,随着视频请求质量的提高,系统产生的延迟大幅度增加.表2展示了使用本机制后直播系统产生的延迟,随着视频请求质量的提高,系统产生的延迟只有小幅度增加.而且,对于流畅型视频请求,使用本机制前后产生的延迟差别不大;但对于高清型和超清型视频请求产生的延迟分别减小了大约1/2和1/3.

4 展望

本文提出了一种基于混合云存储环境的采用可扩展视频编码方式对直播视频进行编码的存储及调度机制.该机制不但考虑了终端用户异构性,而且研究了不同服务器处理性能和不同网络传输性能的差异性对用户的影响,降低了直播系统中的网络延迟,提高了用户端对于直播视频的体验质量.在未来的工作中,本文将实现该机制在基于混合云存储的视频直播系统中

的有效运行,完善本文的工作.

参考文献

- 1 刘志强, 万剑雄, 董秀曼. 云环境视频直播中的服务器选择问题. 计算机工程与设计, 2015, 36(10): 2650-2656, 2661.
- 2 刘景. 基于云计算环境的校园网络视频直播设计. 计算机应用, 2014, 34(2): 572-575, 584.
- 3 孙名松, 赵修娜, 孙西贝, 等. 基于云计算的校园网视频直播系统设计. 哈尔滨理工大学学报, 2012, 17(1): 58-62, 67.
- 4 张欣, 何欣枫, 姜健. 基于应用层组播的实时流媒体负载分配方案. 大众科技, 2011, (10): 42-44. [doi: 10.3969/j.issn.1008-1151.2011.10.017]
- 5 于雷. 校园网网络直播系统设计. 电脑知识与技术, 2014, (10): 2385-2387.
- 6 刘侃, 李俊, 杨坚. 一种共享存储视频服务器集群的请求调度策略. 小型微型计算机系统, 2009, 30(4): 666-670.
- 7 罗建光, 张萌, 赵黎, 等. 基于P2P网络的大规模视频直播系统. 软件学报, 2007, 18(2): 391-399.
- 8 秦丰林, 刘琨. P2P网络流媒体关键技术. 电子学报, 2011, 39(4): 919-927.
- 9 彭宏, 吴海巍, 叶敏展, 等. 基于流媒体的移动视频直播系统的设计与实现. 电子技术应用, 2014, 40(9): 111-113, 117.
- 10 刘奇. P2P视频直播系统改进设计. 现代电子技术, 2011, 34(1): 32-34, 38.
- 11 李长利, 应小昆, 张万光, 等. 基于P2P网络技术的视频直播系统的实现与测试. 全国开放式分布与并行计算学术年会. 扬州, 中国. 2008. 83-86.
- 12 Hajjat M, Shankaranarayanan PN, Maltz D, et al. Dynamic request splitting for interactive cloud applications. IEEE Journal on Selected Areas in Communications, 2013, 31(12): 2722-2737. [doi: 10.1109/JSAC.2013.131212]
- 13 Grandison T, Maximilien EM, Thorpe S, et al. Towards a formal definition of a computing cloud. Proc. of the 2010 6th World Congress on Services. Miami, FL, USA. 2010. 191-192.
- 14 周可, 王桦, 李春花. 云存储技术及其应用. 中兴通讯技术, 2010, 16(4): 24-27.
- 15 刘光, 房胜, 尉希娟. 可伸缩视频编码的码率控制技术. 计算机技术与发展, 2010, 20(10): 26-29.
- 16 Lewis G, Echeverria S, Simanta S, et al. Tactical cloudlets: Moving cloud computing to the edge. Proc. of IEEE Military Communications Conference. Baltimore, MD, USA. 2014. 1440-1446.