

DOI: 10.7652/xjtub201404005

视频点播系统中视频分片协同存储方案研究

赵晓明^{1,2}, 周颢^{1,2}, 何军^{1,2}, 赵保华^{1,2}

(1. 中国科学技术大学计算机科学与技术学院, 230027, 合肥;

2. 安徽省计算与通讯软件重点实验室, 230027, 合肥)

摘要: 为了提高视频点播系统的服务质量,减少等待延迟,改善用户体验,将视频协同存储与网络编码技术相结合,以最大化本地命中为目标,提出了一种视频分片协同存储最大化本地命中算法。利用网络编码技术,预先将视频进行分片编码。在接收到视频请求时,将编码后的视频片发给用户。在客户端接收到足够的视频片后,即可进行解码,从而得到完整的视频。该算法构造了一种资源分配有向图,将原问题转化为该图上的最小费用流问题;通过求解该最小费用流问题,在多项式时间复杂度内得到最优化方案。当视频的总容量与服务器的总容量之比较大时,该算法依然有较好的表现。实验结果表明,该算法的性能优于其他已有的算法,特别是在视频总容量与服务器总容量之比较大的情况下,性能有10%的提高。实验分析了不同切片大小对算法性能的影响,选择了1 MB作为默认的切片大小。

关键词: 网络编码;视频协同存储;最小费用流

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2014)04-0026-05

Research of Collaborative Placement for Segmented Video in VoD System

ZHAO Xiaoming^{1,2}, ZHOU Hao^{1,2}, HE Jun^{1,2}, ZHAO Baohua^{1,2}

(1. School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China;

2. Province Key Laboratory of Software in Computing and Communication, Hefei 230027, China)

Abstract: To improve the quality of service in Video on Demand (VoD) system, reduce delay and improve user experience, by combining network coding with video collaborative placement, the problem of maximum local hits is discussed, and an algorithm named MHPS is proposed. In terms of network coding, the videos are divided into segments in advance, and the clients can decode the video with enough segments. The algorithm constructs a placement graph and converts the original problem as the minimum cost flow over the graph. Thus the optimal solution can be obtained in the polynomial time complexity. The extensive simulations are carried out to evaluate the performance of the proposed algorithm. The results suggest that the proposed algorithm outperforms the other existing solutions with 10% improvement when the total size of videos is close to that of server capacity. The influence of the segment size in the algorithm is also analyzed.

Keywords: network coding; video collaborative placement; minimum cost flow

随着无线网络的迅猛发展,无线网络提供的高速率传输使得视频点播系统 (Video on Demand,

收稿日期: 2013-11-19。 作者简介: 赵晓明(1989—),男,硕士生;周颢(通信作者),男,讲师。 基金项目: 国家自然科学基金资助项目(61371118);中央高校基本科研业务费专项资金资助项目(WK0110000029)。

网络出版时间: 2014-03-04

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20140304.0943.002.html>

VoD)业务呈现爆炸性增长。如何利用有限的资源将海量的视频数据存储起来成为研究热点^[1-4]。

目前,VoD 视频存储研究的主流是利用服务器网络内部的高速链路来整合各个节点的存储能力,从而能够胜任大规模的视频库存储,即协同存储。这需要考虑多方面的因素,例如服务器容量限制、链路带宽限制、视频内容流行程度分布等。本文将视频协同存储与网络编码^[5]相结合,提出了一种视频分块协同存储最大化本地命中算法,能够得到最优化的解决方案,并且当视频的总容量与服务器的总容量之比较大时,依然有较好的表现。

1 背景以及相关工作

1.1 研究背景

无线互联网正在蓬勃发展中,视频流量出现爆炸性增长。VoD 系统由于摆脱了时间、空间的限制,具有方便、快捷的优点,因而具有广泛的应用。如图 1 所示是一种应用场景,利用长期演进(Long Term Evolution, LTE)网络中服务网关上的存储能力为用户提供视频缓存服务。

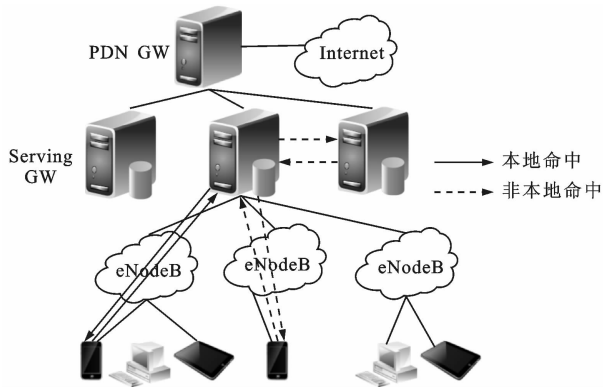


图1 VoD系统在长期演进LTE网络的应用场景

在LTE网络及其他VoD系统应用场景下,为有效存入海量视频库引入了协同存储方案。每个服务器存储视频库的一个子集,并将其存储内容通过直接响应的方式为用户提供服务或者通过协作请求为其他的服务器提供服务。对于用户发出的视频请求,如果对应的服务器上有这个视频的备份,则称之为本地命中;否则,需要从其他的服务器上获取该视频,称之为非本地命中。因此,如何将这些视频分布式协同存储到服务器网络中有重要意义。

1.2 相关工作

在视频协同存储中,对不同的目标进行了研究。Applegate等基于硬盘空间、链路带宽、视频内容流行程度等约束条件,以最小化总的的数据量为目标进

行研究^[1];Borst等将协作存储和路由联系起来,以最大化服务节点能够提供的流量并最小化总带宽的花费为目标进行了探讨^[3]。然而,他们所提算法复杂度太高,应用场景也有局限。Xie等基于底层、非结构化的简单网络,利用交通工程和协作存储从ISP的角度以最小化最大的拥塞程度为目标进行研究^[2]。该算法需要判断请求的模式,这显然不实际。He等提出以最大化系统本地命中为目标的快速近似最优的存储方法,以提高系统本地命中为目标提出了一个快速贪婪算法^[6]。

1.3 基于网络编码的视频协同存储算法

传统的通信网络传送数据的方式是存储转发,即中间节点不对数据内容做任何处理,只进行转发。网络编码^[5](Network Coding, NC)通过中继节点对接收到的信息进行编码来提高多播网络容量。在网络中的各个节点上对各条信道上收到的信息进行线性或非线性的处理,然后转发给下游节点。

在VoD系统中,可以利用网络编码预先将视频进行分片编码。在接收到视频请求时,将编码后的视频片发给用户。客户端接收到足够的视频片后可进行解码,从而得到完整的视频。

采用NC编码能有效提高存储效率,提高系统可靠性。本文研究的目标是将前期的问题^[6]与NC编码的分块存储相结合,提出基于NC编码的视频分块协同存储方案,使得本地命中最大。

2 系统建模

建立一个VoD系统,视频集合为 N ,视频 $n_i \in N$,原始的大小是 s_i 。采用NC编码,每个切片的大小为 $\bar{s}$ 。由NC编码的原理可知,对于视频 n_i ,系统中可以产生任意数目的切片。在这些切片中,用户至少需要获得 $R_i = \lceil s_i / \bar{s} \rceil$ 才能恢复出视频。

服务器集合 M 用来存储视频,服务器 $m_t \in M$ 的可用容量为 W_t 。在实际中, $s_i \ll W_t$ 。假定视频 n_i 在服务器 m_t 上的访问频率为 p_t^i ,视频 n_i 在服务器 m_t 上存放的视频切片块数为 k_t^i ,定义变量如表1所示。

在VoD协同存储中,本地命中越多,网络中流量就会变得越少,从而节省大量的带宽。因此,基于分片的最大化本地命中(Maximum Hit Problem based on Segment, MHPS)问题的表达式为

$$Z = \max \sum_{n_i \in N} \sum_{m_t \in M} p_t^i k_t^i \bar{s} \quad (1)$$

s. t.

$$\sum_{n_i \in N} k_t^i \leq \lfloor W_t / \bar{s} \rfloor \quad m_t \in M \quad (1a)$$

$$k_t^i \leq R_i \quad n_i \in N, m_t \in M \quad (1b)$$

$$\sum_{m_t \in M} k_t^i \geq R_i \quad n_i \in N \quad (1c)$$

约束条件(1a)表示对于服务器 m_t , 所存储的视频的总容量不得大于 m_t 的容量; 约束条件(1b)表示对于视频 n_i , 在服务器 m_t 上放置的片段数目不得大于 R_i ; 约束条件(1c)表示对于视频 n_i , 系统中所存储的切片总数必须能够保证恢复出该视频。

表 1 各种变量的定义列表

符号	含义
n_i	编号为 i 的视频
m_t	编号为 t 的视频
N	视频的集合
M	服务器的集合
s_i	视频 n_i 的大小
$\bar{s}$	每一个视频切片的大小
W_t	服务器 m_t 的容量
p_t^i	视频 n_i 在服务器 m_t 上的访问频率
k_t^i	视频 n_i 存储到服务器 m_t 上的切片个数

3 基于最小费用流问题的 MHPS

最小费用流问题是指在给定网络中求从一个源节点到目的节点流值为某一常数的流, 使其费用最小^[7]。该问题有着广泛的应用, 例如物流管理、供应链研究等领域。

本节构造出一个资源分配有向图 G^{MHPS} , 证明了图 G^{MHPS} 上最小费用流问题的解就是问题(1)的解, 然后给出相应的算法。

3.1 资源分配有向图 G^{MHPS}

定义图 $G^{\text{MHPS}} = (V, E)$, 其顶点集合为 $V = M \cup N \cup \{s, t\} \cup \{s_r, v_r\}$, 其中 s_r, v_r 为虚拟节点, $|V| = O(N + M)$ 。

图 G^{MHPS} 的边主要由 7 类组成, 如表 2 所示。

表 2 G^{MHPS} 边的类型

类型	含义
E_1	$\{ \langle s, m_t \rangle \mid m_t \in M \}$
E_2	$\{ \langle m_t, n_i \rangle \mid n_i \in N, m_t \in M \}$
E_3	$\{ \langle n_i, t \rangle \mid n_i \in N \}$
E_4	$\{ \langle n_i, v_r \rangle \mid n_i \in N \}$
E_5	$\{ \langle m_t, s_r \rangle \mid m_t \in M \}$
E_6	$\{ \langle s_r, v_r \rangle \}$
E_7	$\{ \langle v_r, t \rangle \}$

假设边 e 的起始顶点为 e_{start} , 终止顶点为 e_{end} , 边 e 的容量函数为

$$\text{cap}(e) = \begin{cases} \lfloor W_t / \bar{s} \rfloor, & e \in E_1 \& e_{\text{end}} = m_t \\ R_i, & e \in E_2 \& e_{\text{end}} = n_i \\ R_i, & e \in E_3 \& e_{\text{start}} = n_i \\ \infty, & e \in E_4 \cup E_5 \cup E_6 \\ \sum_{m_t \in M} \lfloor W_t / \bar{s} \rfloor - \sum_{n_i \in N} R_i, & e \in E_7 \end{cases}$$

边 e 的花费函数为

$$\text{cost}(e) = \begin{cases} -p_t^i, & e \in E_3 \& e_{\text{end}} = n_i \& e_{\text{start}} = m_t \\ 0, & e \notin E_3 \end{cases}$$

当 $|M| = 2, |N| = 3$ 时, G^{MHPS} 系统如图 2 所示。

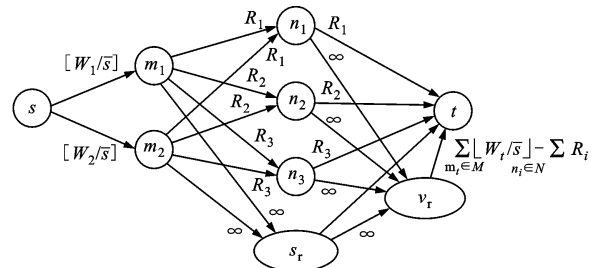


图 2 $|M| = 2, |N| = 3$ 时的 G^{MHPS} 图

在 G^{MHPS} 中, 从 s 节点流入的流量即为系统中服务器的总容量。 s 到 m_t 边的容量为服务器 m_t 所能存放的切片数, 即 $\lfloor W_t / \bar{s} \rfloor$ 。由于从 s 流入的流量为所有 s 到 m_t 的边容量之和, 所以每个 s 到 m_t 边的流量也一定是 m_t 能够存放的切片数目。

边 m_t 到 n_i 的容量是恢复 n_i 所需要最小切片数, 而其上的流量则表示 n_i 在 m_t 上放置的片数。边 n_i 到 t 的容量是恢复 n_i 所需要最小切片数, 由网络流的流量守恒特性以及 v_r 到 t 的容量可知, 边 n_i 到 t 的流量等于边的容量。

此时, 视频 n_i 在服务器 m_t 上的存放片数 k_t^i 即为从服务器 m_t 到该视频 n_i 的流量 (G^{MHPS} 图中网络流的单位为切片的大小), 从而求解出 MHPS 问题。例如, 图 3 表示 G^{MHPS} 的一个实例, 边上的数字表示分配的流量, 其中视频 n_1 从服务器 m_1 和 m_2 上分别得到流量为 4 和 3, 则对应的分配方案就是 $k_1^1 = 4, k_2^1 = 3$, 其他同理。

因此, 若图 G^{MHPS} 上可行解为 ϕ , 边 $\langle m_t, n_i \rangle$ 的流量为 MHPS 中 k_t^i 的值, 由 k_t^i 组成的解决方案为 MHPS 的解。

3.2 G^{MHPS} 最小费用流的解为 MHPS 问题的解

设 ϕ 是 G^{MHPS} 的一个解, ξ 是 MHPS 的一个解。

下面从两个方面来证明 ψ 与 ξ 一一对应。

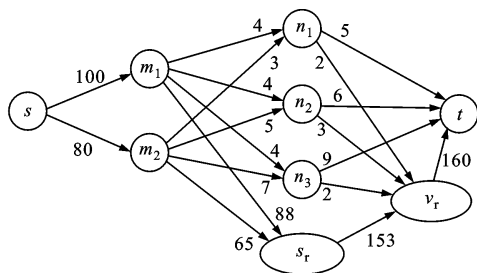


图3 图 G^{MHPS} 与分配方案 k_i^j 的关系

引理 1 G^{MHPS} 的解 ψ 是 MHPS 的解 ξ 。

假定 ψ 是 G^{MHPS} 的一个解, 根据 G^{MHPS} 中 E_1 集合的容量函数定义可知, 服务器 m_i 的流量最大为 $\lfloor W_i/\bar{s} \rfloor$, 满足约束条件(1a); 同时, 根据 G^{MHPS} 中 E_2 集合容量函数定义可知, 视频 n_i 放在服务器 m_i 上的数目不超过 R_i , 满足约束条件(1b); 由 v_i 到 t 的容量限制可知, 每一个视频 n_i 到 t 的流量至少为 R_i , 满足约束条件(1c)。因此, ψ 中服务器到视频集合的分配方案是 MHPS 的一个解 ξ 。

引理 2 MHPS 的解 ξ 是 G^{MHPS} 的解 ψ 。

假定 ξ 是 MHPS 的一个解, 根据 ξ 中视频到服务器的分配方案, 依照 G^{MHPS} 的定义, 可以构建出一个对应的图 G^{MHPS} 的一个解 ψ 。因此, 由 ξ 构造的解 ψ 是 G^{MHPS} 的一个解。

综上, G^{MHPS} 的解与 MHPS 的解一一对应。因此, 以下定理 1 显然得证。

定理 1 G^{MHPS} 上最小费用流问题的解是 MHPS 问题的最优解。

3.3 算法及其时间复杂度

最小费用流问题已经有较多的实现^[8-9], Orlin 等提出了多项式时间的原始网络单纯形算法^[8]。在仿真实验中, 利用 LEMON^[10] 函数库进行求解。

G^{MHPS} 中, $|V| = O(N + M)$, $|E| = O(NM)$, G^{MHPS} 的时间复杂度为 $O(NM)$ 。LEMON 库基于网络单纯形算法, 其复杂度是 $O(t^2 s^2 \log t)$ ^[8], 其中 t 和 s 分别是网路中节点数和边数。因此, 本算法时间复杂度是 $O((NM)^2 (N + M)^2 \log(N + M))$ 。在实际系统中, $|N| \gg |M|$, 所以本算法的时间复杂度为 $O(N^4 M^2 \log N)$ 。

4 实验仿真

4.1 仿真环境

在仿真实验中, 建立一个 VoD 缓存系统, 其中单个服务器的容量大小从 120 GB 到 240 GB 随机

分布, 单个视频大小从 20 MB 到 400 MB 随机分布。

在实验中, p_i^j 取决于视频 n_i 的流程度和服务服务器 m_i 的用户数量。用户数量为 20 到 30 的随机整数。在每一个服务器上, 所有视频的分布为 Zipf 分布, 指数在 0.7~0.9 之间随机选择。为了模拟视频分布的多样性, 视频在每一个服务器上的流程度随机产生。访问频率 p_i^j 就是视频 n_i 在服务器 m_i 上的 Zipf 因子与服务器 m_i 的用户数量乘积。

4.2 性能比较

这里将 MHPS 算法和两种算法进行了比较。第一种是文献[6]中提出的 MHP 算法, 不考虑视频切片问题, 直接将完整的视频存储于某个服务器。第二种算法是贪心算法 MHPSG, 即以访问频率的大小顺序来遍历视频库, 将其存放到各个最优的服务器上, 从而保证系统中存入了完整的视频库, 运用多重背包问题将服务器剩余的空间尽可能多存储。

图 4 绘制了在服务器数目分别为 10 和 20 的情况下, 当视频数目从 100 变到 1 600 时 3 种算法的性能。由图 4 可见: ①当服务器的数目和容量保持不变的时候, 系统所获的收益呈现下降趋势, 这是由于大量的低价值的视频分片占据了系统的容量, 当视频片段数目达到一定的数目时, 由于视频片段的总空间大于服务器的总空间, 因而出现无解的情况; ②对比不同的算法可以看出, MHPS 算法优于 MHP 和 MHPSG 算法, 随着视频总容量的增加, MHPS 算法的优势更加明显, 例如当视频总容量达到服务器总容量的 90% 左右时, MHPS 算法的性能超过 MHP 算法 10% 以上。

4.3 $\bar{s}$ 的取值对 MHPS 算法的影响

对于 MHPS 算法, 显然切片大小对系统的性能有影响。图 5 中绘制了在不同系统规模下, 切片大小对系统性能的影响。

当切片数目过小时, 由于大量收益低的视频切片会降低整个视频库视频切片的平均收益, 从而导致系统性能下降。当切片数目过大时, 容易造成系统空间浪费, 导致系统性能下降。由图 5 可见, 当切片大小为 1 MB, 系统具有最佳的性能。因此, 本文中选择了 1 MB 作为默认的切片大小。

5 结论

VoD 视频协同存储对于提高服务质量、提升用户体验、减小等待延迟等具有重要的意义, 本文通过基于 NC 编码, 对 VoD 视频协同存储问题进行了探究, 提出了一个 VoD 系统中基于分片的视频协同存

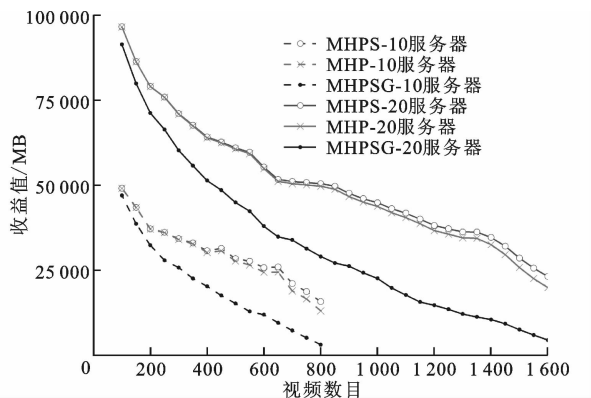


图4 不同算法获得的收益比较

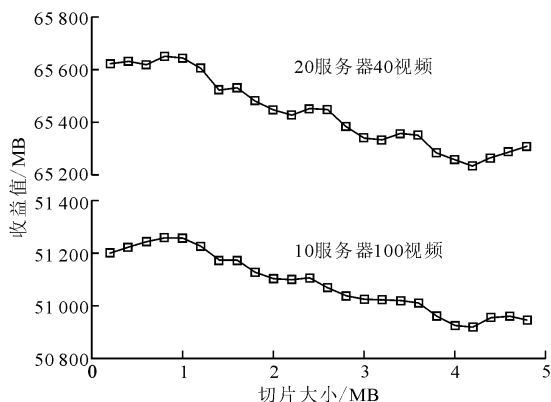


图5 不同系统规模下切片大小对系统性能的影响

储解决方案 MHPS, 该算法具有多项式时间复杂度。通过仿真, 可以看出 MHPS 算法能够实现最优化的收益, 该算法优于 MHP 算法, 特别是在视频库的容量接近服务器总容量的情况下, MHPS 算法能够提供较大的收益提升。

参考文献:

- [1] APPLEGATE D, ARCHER A, GOPALAKRISHNAN V, et al. Optimal content placement for a large-scale VoD system [C]//Proceedings of the 6th International Conference on Emerging Networking Experiments and Technologies. New York, USA: ACM, 2010: 4.
- [2] XIE Haiyong, SHI Guangyu, WANG Pengwei. TECC: towards collaborative in-network caching guided by traffic engineering [C]//Proceedings of the International Conference on Computer Communications. Piscataway, NJ, USA: IEEE, 2012: 2546-2550.
- [3] BORST S, GUPTA V, WALID A. Distributed caching algorithms for content distribution networks [C]//Proceedings of the International Conference on Computer Communications. Piscataway, NJ, USA: IEEE, 2010: 1-9.

- [4] ADAMIC L A. Zipf, power-laws, and pareto: a ranking tutorial [EB/OL]. (2000-04-10) [2013-11-16]. <http://impact.asu.edu/cse591sp11/ZipfParetoPower-ranking.pdf>.
- [5] AHLWEDE R, CAI Ning, LI S Y R, et al. Network information flow [J]. IEEE Transactions on Information Theory, 2000, 46(4): 1204-1216.
- [6] HE Jun, ZHAO Xiaoming, ZHAO Baohua. A fast, simple and near-optimal content placement scheme for a large-scale VoD system [C]//Proceedings of the 2012 IEEE International Conference on Communication Systems. Piscataway, NJ, USA: IEEE, 2012: 378-382.
- [7] AHUJA R K, MAGNANTI T L, ORLIN J B. Network flows: theory, algorithms, and applications [J]. Journal of the Operational Research Society, 1994, 45(11): 1340-1340.
- [8] ORLIN J B. A polynomial time primal network simplex algorithm for minimum cost flows [J]. Mathematical Programming, 1997, 78(2): 109-129.
- [9] DANTZIG G B. Linear programming and extensions [M]. Berlin, Germany: Springer-Verlag, 1998: 404-412.
- [10] Egerváry Research Group on Combinatorial Optimization. Library for efficient modeling and optimization in networks [EB/OL]. (2011-06-07) [2013-11-16]. <http://lemon.cs.elte.hu>.

[本刊相关文献链接]

- 李巍, 李国兵, 朱世华, 等. 一种采用空时网络编码的多向中继通信方案. 2014, 48(2): 80-85. [doi:10.7652/xjtuxb201402014]
- 崔华力, 钱德沛, 张兴军, 等. 用于无线多跳网络视频流传输的优先级机会网络编码. 2013, 47(12): 13-18. [doi:10.7652/xjtuxb201312003]
- 吕政, 余志军, 刘海涛, 等. 协作通信中联合信道-网络编码的性能分析与资源分配. 2012, 46(4): 83-87. [doi:10.7652/xjtuxb201204014]
- 池信泽, 周颖, 赵保华. 面向前向纠错的无线 Mesh 网多播差错控制协议. 2011, 45(8): 30-36. [doi:10.7652/xjtuxb201108006]
- 卢冀, 肖嵩, 吴成柯, 等. 无线网络中应用机会式网络编码的广播重传方法. 2011, 45(2): 68-72. [doi:10.7652/xjtuxb201102014]
- 全杰, 钱德沛, 刘轶, 等. 针对事件驱动型传感器网络的多路径编码路由协议. 2010, 44(6): 39-45. [doi:10.7652/xjtuxb201006008]

(编辑 武红江)