

# 基于云存储的 P2P 视频点播系统数据复制策略

王娟

(江苏宿迁泽达职业技术学院, 江苏 宿迁 223800)

**摘要:** 交互性支持对 P2P 视频点播系统具有重要的意义, 视频点播服务的大规模普及离不开用户交互性的支持。讨论了如何有效利用对等节点的带宽和存储资源来主动复制数据块, 提出了一种基于云存储的数据复制策略 CSPR。仿真实验结果表明, 相比于现有的数据复制策略, CSPR 可以显著提高用户进行随机搜索操作时的响应速度, 并降低网络复制开销。

**关键词:** 对等网络; 视频点播; 云计算; 数据复制

中图分类号: TP393

文献标识码: A

文章编号: 1674-7720(2013)20-0070-03

## The strategy based on the Cloud storage data replication for Peer-to-Peer video-on-demand systems

Wang Juan

(Suqian Zeda Vocational & Technical College, Suqian 223800, China)

**Abstract:** It is important for P2P VoD systems to support user interaction. And it is inseparable for VoD service mass popularity to support user interaction. This paper discusses how to effectively use the bandwidth and storage resources of peer nodes to actively replication blocks of data. And it proposes a strategy called CSPR which is based on the Cloud storage data replication. The simulation results show that, compared with the existing data replication strategy, CSPR can significantly enhance the response of user random search operations, and reduce the overhead of replications.

**Key words:** P2P; video-on-demand; Cloud computing; data replication

视频点播 VoD (Video on Demand) 服务是指根据用户的需要, 随机选择视频内容进行播放的服务。在 P2P 网络上构建视频点播系统能有效利用节点间网络带宽、存储空间和数据资源, 系统具有良好的可扩展性、数据可用性和可靠性。因此, P2P 视频点播系统 (P2P VoD) 已成为当前一个重要的研究领域。但 P2P VoD 中每个节点随时都可能暂时或永久离开系统, 这使得构建支持交互性的 P2P VoD 极富挑战性。系统中的节点均负责存储数据, 一旦某节点暂时离开, 存在其上的数据就暂时不可访问, 而节点的永久离开更会造成数据的丢失。因而采用有效的办法来应对动态的系统环境, 从而满足用户的随机搜索操作 (VCR) 则显得十分必要。

云计算是当今 IT 界的热门技术, 借助云计算, 网络服务提供者可以在瞬息之间处理数以千万计甚至亿计的信息, 实现与超级计算机同样强大的效能。当云计算系统运算和处理的核是大量数据的存储和管理时, 云计算系统中就需要配置大量的存储设备, 此时云计算系

统就转变成为一个云存储系统, 所以云存储是一个以数据存储和管理为核心的云计算系统。

在 P2P VoD 中, 资源保存在用户的机器上, 视频文件越流行, 下载的用户越多, 资源的副本越多, 下载效率就越高。但对于一些过时的文件, 用户会选择删除, 以免占用过多的个人空间。因此, 一些冷门文件就不容易通过对等节点间的共享获得。参考文献 [1] 指出用户访问的集中度符合“二八定律”, 即 20% 的视频内容吸引了 80% 的用户访问, 而其他的 80% 的用户访问只吸引了 20% 的用户访问量。在基于云存储的 P2P 视频点播系统中, 考虑将存储空间当作一片“云”, 在无限量的空间, 当用户想从个人的机器上删除一些过时但非毫无价值的的数据块时, 可以选择把资源通过数据复制的方式放在云上, 从而把数据块保存起来。这样用户不仅可以通过节点间的共享快速获得热门视频数据, 也可以通过检索找到被存储起来的冷门视频数据, 从而可以快速地响应用户的 VCR 操作。

欢迎网上投稿 [www.pcachina.com](http://www.pcachina.com) 73

## 技术与方法 Technique and Method

本文主要讨论在节点将要离开系统时,根据视频数据各个部分的受欢迎程度以及对等节点间的网络距离,提出了一种基于云存储的数据复制策略 CSPR。仿真实验结果表明,相比于现有的基于流行度的数据复制策略 PPR, CSPR 可以显著提高用户执行 VCR 操作的响应速度,并有效降低复制数据时所产生的网络开销。

### 1 相关工作

VCR 功能是对等点播系统中最重要的功能之一,也是本文研究的主要目标。在树状 P2P VoD 系统中,如 P2Cast<sup>[2]</sup>则没有考虑 VCR 操作对系统性能造成的影响,没有针对 VCR 操作设计新的数据存储策略。当用户进行 VCR 操作时,节点需要通过重新加入组播树。

在网状 P2P VoD 系统中,如 VMesh<sup>[3]</sup>和 BulletMedia<sup>[4]</sup>则是利用 DHT 技术支持 VCR 操作。VMesh 通过本地向下/向上的指针和 DHT 查询来支持交互式操作。在 VMesh 中,节点通过本地存储的数据块来为其他节点提供服务,而不是正在收看的数据块。存储的数据块在节点的上线时间内不变。BulletMedia 通过 DHT 方法让对等节点大公无私地主动查找稀有的数据块并进行复制存储,保证每个数据块在系统中至少有  $K$  个副本。当用户进行 VCR 操作时,通过 DHT 方法查找相应的数据片段,以减轻用户进行 VCR 操作时对服务器造成的压力。

针对因节点离开而导致数据丢失的问题,参考文献[5]采用了基于预测的带宽分配策略 PBA。该策略的核心思想是让父节点给那些具有较高带宽的子节点提供较大的上传带宽,并且它仅仅让节点把数据复制给将来可能需要的那些节点。但 PBA 策略是假定用户从头到尾观看视频节目的,并且在观看过程中没有 VCR 操作,这是不符合视频点播系统中的用户观看行为的,因此 PBA 策略不能直接应用于 P2P VoD 系统中。

研究表明,用户在视频点播系统中的行为并非毫无规律。参考文献[6]指出,不同的文件具有不同的流行度,即用户访问的频率;大部分的用户访问都集中于少量热门节目上,并且文件的流行度是随时间而改变的。

### 2 基于云存储的数据复制策略

#### 2.1 云节点的选择策略

为了降低节点离开对 P2P VoD 系统性能的影响,常根据特定的选择策略从系统的可用节点集中选择部分节点来使用<sup>[7]</sup>。节点选择策略可分为随机选择和确定性选择两大类,其中随机选择策略是指不考虑节点的特征随机选择节点;确定性选择策略是指根据节点的某种特征选择满足一定标准的节点,如根据节点的当前在线时长、带宽吞吐能力等选择节点<sup>[8-9]</sup>。

由于节点的当前在线时长对其余留时长有一定的预见性,因此选择最长当前在线时长的策略用于 DHT 网络中邻居节点的选择<sup>[10]</sup>和超级节点的选择<sup>[11]</sup>,以及覆盖层多播树种双亲节点的选择<sup>[12]</sup>。GODFREY B P 等人<sup>[7]</sup>将不同的选择策略应用于 5 个当前广泛使用的对等网

络中。实验结果表明,与其他选择策略相比,随机选择策略可以有效降低节点动态性对 P2P 网络的影响。这是由于随机选择策略具有简单、易实现等特点。因而随机选择策略被广泛地应用于 P2P 网络中,如 P2P 文件存储系统 Total Recall<sup>[13]</sup>等。

因此,本文采用随机选择策略来选择云存储中的云节点。

#### 2.2 云节点的组织方式

基于 P2P 的系统中两节点间网络距离的测算可通过计算系统中数据传输时延、网络最小带宽和数据传输所经路由跳数等方法获得。对于 P2P 视频点播系统中数据块的可用性和可靠性来说,数据传输时延是一个很重要的因素,而数据传输时延最主要的影响因子是传输两节点间的可用带宽。

P2P 覆盖网技术从最初以 Napster 为代表的有着中央目录服务器的对等网络结构,到后来以 Gnutella 为代表的完全分布式的对等网络和提供匿名发布和获取文档的 Freenet,再到以 CAN、Chord、Pastry 和 Tapestry 等为代表的基于分布式哈希表的结构化对等网络,对等网络的发展大致经历了 3 个阶段,每个阶段都采用了不同的资源定位和路由模型。根据拓扑结构关系可以将 P2P 系统分为 4 类:集中式 P2P 系统、完全分布式 P2P 系统、混合式 P2P 系统和结构化 P2P 系统。

本文将 P2P 视频点播系统中网络距离相近的节点划分到同一云中,以实现节点的分组和有效管理,采用参考文献[14]中的节点分组算法来对云节点进行组织。该算法的核心思想是依次从初始化网络中随机选取一个节点,以该节点为圆心,预测网络距离为半径,其内的所有节点都划入一个分组,再将获得分组的节点从初始网络中移除。以此方法循环划分,可将初始网络中大部分节点划入分组。

另外,本文采用参考文献[15]中的方法对分布式数据块的流行度进行估计。

#### 2.3 基于云存储的数据复制算法

基于云存储的数据复制策略包括以下 4 个步骤:(1)根据随机选择策略选择云节点;(2)按照网络距离分组算法组织这些数据复制的目标云节点;(3)根据分布式平均算法估计出每个数据块的流行度,并按从高到低的顺序排序,按照用户存储空间的大小确定数据块集合  $S$ ;(4)把  $S$  以外的数据块推向云节点中进行存储。该算法可用伪代码表示如下:

```
for( $1 \leq m \leq t$ )    compute( $\hat{p}_m$ );           //计算数据块的流行度
if(delete(M))
{
for( $1 \leq m \leq t$ )
{
NodeGet( $n, i$ );           //从当前在线的节点中随机选取
                           节点  $i$  作为数据复制目标云节点
```

## 技术与方法 Technique and Method

```
SegmentChoose(P, M, m); //选取流行度高且副本数
                          //低的数据块 m
Replication(m, i); //将数据块 m 复制给节点 i
}
```

代码中,  $n$  表示在线的节点数;  $i$  为选择的目标数据复制云节点;  $m$  表示选择的数据块;  $t$  表示将要删除数据块的个数;  $T$  表示整个数据块的集合;  $S$  为各个节点的性能值的集合, 其中  $s_1 \geq s_2 \geq \dots \geq s_n$ ;  $M$  表示节点将要删除的数据块的集合,  $M=T-S$ ;  $P$  表示各个数据块的流行度集合, 其中  $\hat{p}_1 \geq \hat{p}_2 \geq \dots \geq \hat{p}_t$ 。

### 3 仿真实验

#### 3.1 关键性能指标

选择了如下两个重要参数作为 P2P 视频点播系统中数据复制算法性能的关键指标, 也是仿真实验的测量对象。

(1) 响应用户 VCR 操作时延: 数据复制的目标之一是快速响应用户进行 VCR 操作的时延, 以改进用户的播放体验。该指标指从用户进行 VCR 操作开始到获得所想要的数据块为止所需要的时间, 响应用户 VCR 操作时延的减少表示数据复制算法的性能的提高。

(2) 网络开销: 由复制的数据块总量来表示, 主要衡量数据复制导致的系统资源的额外消耗, 例如, 增加的对等节点之间网络通信流量。

#### 3.2 实验与结果

仿真实验对本文提出的基于云存储的 P2P 视频点播系统数据复制策略 CSPR 与现有的数据复制策略的关键性能指标进行对比。基于流行度的数据复制策略 (PPR) 在进行数据复制时仅考虑每个数据块的流行度, 而并未考虑节点间的网络距离。

利用 P2PStrmsim 作为仿真工具对本文提出的基于云存储的数据复制策略的性能进行详细评估。在仿真实验中, 使用强度为 20 的 Poisson 过程来模拟节点加入过程, 即平均每秒有 20 个节点进入系统。根据参考文献 [16], 对等网络中的节点在线时间分布可以用 Weibull( $\lambda, k$ ) 分布 (概率密度函数  $f(x)=1-e^{-(x/\lambda)^k}$ ) 来模拟, 取  $\lambda=600$ ,  $k=2$ , 则节点在线时间约为 208 s。

从图 1 可以看出, 随着点播系统中对等节点数目的增加, CSPR 和 PPR 响应用户 VCR 操作时延都呈现增长

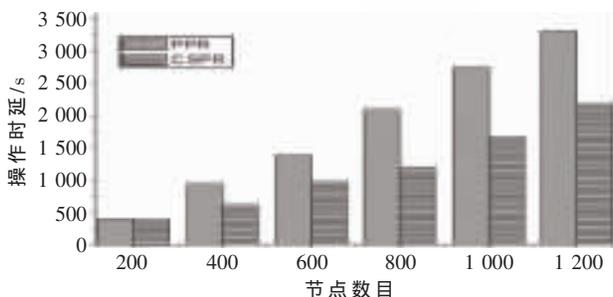


图 1 节点数目和响应用户 VCR 操作时延的关系

的趋势, 但相比于 PPR 策略, CSPR 策略响应用户 VCR 操作的时延明显减小, 因而可以更快速地响应用户的 VCR 操作, 从而改善用户的点播体验。

图 2 所示为复制所产生的开销与仿真时间的关系, 本文使用数据块的数目表示复制开销。从图中可以看出, 随着仿真时间的推移, 两种复制策略所产生的复制开销都在增大, 这主要是由于有更多的用户加入点播系统。但在相同的仿真时间下, 基于云存储的数据复制策略 CSPR 比基于流行度策略 PPR 所复制的数据块少, 这主要是因为 CSPR 策略在进行数据复制时, 不仅考虑了对等节点间的网络距离, 而且考虑了那些不流行的数据块的复制存储。

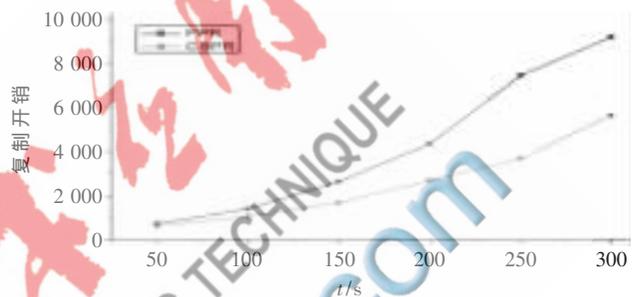


图 2 复制开销与仿真时间的关系

本文在对 P2P 对等点播系统用户动态性分析的基础上, 提出了一种基于云存储的数据复制策略 CSPR。仿真实验结果表明, 相比单纯基于数据流行度的数据复制机制, 本文提出的数据复制机制可较快地响应用户 VCR 操作, 同时降低了网络开销。在后续的工作中, 将考虑对 P2P 视频点播系统的移动用户行为特征进行分析, 将之应用于数据复制策略中, 来更好地满足移动用户观看视频文件时的交互式操作, 从而改善用户的播放体验。

#### 参考文献

- [1] Yu Hongliang, Zheng Dongdong, ZHAO B Y, et al. Understanding user behavior in large-scale video-on-demand systems[C]. In Proceedings of ACM, Eurosys, 2006.
- [2] GUO Y, SUH K, KUROSE J, et al. P2Cast: peer-to-peer patching scheme for VoD service[C]. In Proceedings of the 12th International World Wide Web Conference, Budapest, 2003: 301-309.
- [3] YIU K, JIN X, CHAN G. VMesh: distributed segment storage for peer-to-peer interactive video streaming[J]. IEEE Journal on Selected Areas in Communications(JSAC), 2007, 25(9): 1717-1731.
- [4] NEVENA V, PRIYA G, NIKOLA K, et al. Enabling DVD-like features in P2P video-on-demand systems[C]. In Proceedings of the SIGCOMM Peer-to-Peer Streaming and IP-TV Workshop, Kyoto, 2007: 329-334.
- [5] CHENA Z J, XUE K P, HONG P L, et al. Differentiated bandwidth allocation for reducing server load in P2P VoD[C]. In Proceedings of the 8th International Conference on Grid

- and Cooperative Computing, Lanzhou, 2009: 31-36.
- [6] Luo Jianguang, Tang Yun, Zhang Meng, et al. Characterizing user behavior model to evaluate hard cache in peer-to-peer based video-on-demand service[C]. In Proc. MMM, Kyoto, 2005: 125-134.
- [7] GODFREY B P, SHENKER S, STOICA I. Minimizing churn in distributed systems[C]. In Proc. of the ACM SIGCOMM 2006, New York: ACM Press, 2006: 147-158.
- [8] 张宇翔, 杨冬, 张宏科. P2P 网络中 Churn 问题研究[J]. 软件学报, 2009, 20(5): 1362-1367.
- [9] RHEA S, GEELS D, ROSCOE T, et al. Handling churn in a DHT[C]. In Proc. of the USENIX Annual Technical Conference. Boston: USENIX Association Press, 2004: 127-140.
- [10] LEDLIE J, SHNEIDMAN J, AMIS M, et al. Reliability and capacity-based selection in distributed hash tables[R]. Technical Report, Harvard University Computer Science, 2003: 1-14.
- [11] GARCES E L, BIRSACK E W, ROSS K W, et al. Hierarchical peer-to-peer systems[C]. In Proc. of the ACM/IFIP Int'l Conf. on Parallel and Distributed Computing (Euro-Par). Berlin: Springer-Verlag, 2003: 1230-1239.
- [12] SPRIPANIDKULCHAI K, GANJAM A, MAGGS B, et al. The feasibility of supporting large-scale live streaming applications with dynamic application end-points[C]. In Proc. of the ACM SIGCOMM. New York: ACM Press, 2004: 107-120.
- [13] BHAGWAN R, TATI K, CHENG Y, et al. Total recall: system support for automated availability management[C]. In Proc. of the 1st ACM/Usenix Symp. on Networked Systems Design and Implementation (NSDI 2004). San Francisco: USENIX Press, 2004: 337-350.
- [14] 杨磊, 黄浩, 李仁发, 等. 一种基于分组管理的混合式 P2P 存储系统[J]. 计算机科学, 2010, 37(1): 64-67.
- [15] MEHYAR M, SPANOS D, PONGSAJAPAN J, et al. Asynchronous distributed averaging on communication networks[J]. IEEE/ACM Transactions on Networking, 2007, 15(3): 512-520.
- [16] The Network Simulator-ns-2[EB/OL]. [2013-05-25]. <http://www.isi.edu/nsnam/ns>.

(收稿日期: 2013-06-25)

作者简介:

王娟, 女, 1983 年生, 硕士, 主要研究方向: 计算机网络与多媒体技术。