

P2P 缓存系统中总开销最小的协作缓存策略

刘银龙¹, 汪敏², 马伟¹, 周旭¹, 胡亚辉¹

(1. 中国科学院 信息工程研究所, 北京 100093; 2. 北京城市学院 信息学部, 北京 100083)

摘要: 为降低 P2P 缓存系统中的全局开销, 提出一种基于总开销最小的协作缓存策略。该策略综合考虑 P2P 缓存系统中的传输开销和存储开销, 使用跨 ISP 域间链路开销、流行度、文件大小、存储开销来衡量文件的缓存增益。需要替换时, 首先替换掉缓存增益最小的内容。实验结果表明, 所提策略能够有效降低系统的总开销。

关键词: P2P 缓存; 协同缓存; 总开销; 缓存增益

中图分类号: TP393

文献标识码: A

Cooperative caching scheme based on the minimization of total cost for P2P caches

LIU Yin-long¹, WANG Min², MA Wei¹, ZHOU Xu¹, HU Ya-hui¹

(1. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China;

2. School of Information, Beijing City University, Beijing 100083, China)

Abstract: To reduce the total cost of P2P cache system, a cooperative cache scheme based on the minimization of total cost is proposed. In the scheme, delivery cost and storage cost are taken into account, and inter-ISP cost, popularity, file size, storage cost are used to evaluate each object's caching gain value, which is a new concept defined to estimate the benefits of storing or replacing an object. When a replacement is needed, the objects with the minimum caching gain value will be evicted. Simulation results show that the proposed scheme can effectively reduce the total cost of P2P cache system.

Key words: P2P cache; cooperative cache; total cost; caching gain

1 引言

P2P 缓存 (P2P cache)^[1,2]的基本思想是在网络边缘部署存储设备以将用户访问或下载的热点内容在存储设备中保存一个副本, 当内容再次被用户访问或下载时, 由 cache 中保留的副本直接提供, 以达到加速访问或下载、降低网间带宽和依赖的目的。其特点是“一次外网访问, 多次内网服务”通过部署 P2P 缓存系统, 网络运营商可以利用有限的投入支撑不断增长的 P2P 业务需求, 缓解 P2P 应用带来的带宽冲击。另外, P2P 缓存系统的使用还可

以改善 P2P 应用的用户体验, 为基础网络运营商带来更多的客户。由此可见, P2P 缓存系统的部署不仅可以缓和 P2P 技术和基础网络运营商的矛盾, 而且可以引导二者协同发展, 创造出双赢局面。因此, P2P 缓存系统在现网中得到了广泛使用。

虽然 P2P 缓存系统结构能够完成基本的缓存处理操作, 但要使 P2P 缓存系统应用到现实网络中, 还需要在缓存效率、版权保护以及系统部署等多方面进行优化。本文主要对 P2P 缓存系统中的缓存优化问题进行研究。

P2P 缓存系统又分为单独缓存和协作缓存 2 种。

收稿日期: 2013-11-28; **修回日期:** 2014-8-29

基金项目: 国家自然科学基金资助项目 (61303251, 61102076); 国家科技重大专项基金资助项目 (2013ZX03002-001); 中国科学院战略性先导科技专项基金资助项目 (XDA06010302)

Foundation Items: The National Natural Science Foundation of China (61303251, 61102076); The National Science and Technology Major Project (2013ZX03002-001); The Strategic Pilot Project of Chinese Academy of Sciences (XDA06010302)

单独缓存是针对单个域(可以是一个城域网或省域网等区域网络)进行独立部署,各个域的缓存节点与其他域之间没有联系,不能为其他域的用户提供服务。协作缓存指各个域所部署的缓存节点可以互相协作,共享所缓存的资源并为全局所有用户提供服务,可以提高整个缓存系统的效率。当前的缓存策略都不考虑缓存节点缓存数据的存储开销,只是根据系统的传输开销来决定要缓存或替换的文件。

本文通过综合考虑缓存开销、传输开销,将内容放置在使系统总开销最小的缓存节点,并在缓存替换时,替换掉缓存增益最小的内容。该策略的基本思想包括:1)采用传输开销和存储开销来衡量文件的缓存增益,判断是否值得缓存;2)缓存节点需要缓存新文件时,如果剩余缓存空间不足,将优先替换缓存节点中缓存增益最小的文件。所提策略的目标是降低运营商部署 P2P 缓存系统的总开销。

2 相关研究现状

目前,研究人员已针对 P2P 缓存系统提出了多种缓存策略,主要采用缓存对象的属性来评价对象(或文件,本文将交替使用对象或文件来表示所缓存的内容)的价值,当缓存空间不足时,优先替换掉那些价值最小的对象。

2.1 独立缓存策略

在独立缓存系统中,缓存的收益主要由缓存对象在本域内的流行程度来决定,因此独立缓存策略的目标是根据对象的属性查找并缓存本域内流行程度最高的文件。已有的独立缓存策略主要有以下几种^[3-5]。

最小相对大小优先替换策略。根据各个文件的存储比例,当缓存空间不足时,优先替换掉缓存空间中存储百分比最少的文件。

发送字节数最少优先替换策略。根据每个已缓存的文件为用户服务所发送的字节数,当缓存空间不足时,优先替换那些为用户发送的字节数据字节数最少的文件。

P2P 缓存策略。将 P2P 文件分成较小的段,并根据用户对文件请求量的增加缓存该文件的更多分段。这样,热度较低的文件最多只会被缓存几个分片,节省了大量的缓存空间,提高了缓存空间的利用率,同时,也避免了缓存内容的频繁换入换出。

但是,单个缓存节点的缓存能力有限,各域的缓存节点仅能存储少量的内容,很多用户仍然需要

从外网下载所需要的文件。

2.2 协作缓存策略

在协作 P2P 缓存技术的研究中,需要解决的一个关键的问题是如何合理地调度和管理 P2P 缓存系统中的缓存资源,这依赖于缓存系统所采用的协作缓存策略。协作缓存系统根据缓存策略决定在各个存储节点缓存哪些文件,并选择最佳的存储节点为用户服务。

针对 P2P 缓存系统的协作,文献[6]提出了 2 类策略。第一类是自私的协作缓存策略,各域的缓存节点可以从其他域的缓存节点获取文件,但是各缓存节点仍然根据文件在本域的流行程度决定存储哪些文件。由于文件在各域的流行程度很接近,大部分文件会被存储在多个缓存节点,这类策略会导致协作缓存系统中所存储的文件有很大的冗余。第二类是协同的协作缓存策略,各域的缓存节点互相协作并根据文件在全局的流行程度存储文件。如果一个域的缓存节点已经存储了某个文件,其他域的缓存节点就不会再存储该文件,这些域的用户都需要从已存储该文件的缓存节点获取该文件。由于在域间传输数据也会耗费一定的链路代价,对于比较热门的文件,频繁地从其他域的缓存节点获取也会产生比较大的开销。

上述协作策略都是针对同一运营商的不同域间提出来的,Dai 和 Dan 分别在文献[7]和文献[8]中对部署在同一地区不同运营商的网络中的 P2P 缓存系统之间的协作进行了研究,并在假设运营商之间可以免费传输流量(这些运营商称为对等运营商)的情况下,提出了 2 种协作缓存策略。第一种是自私的协作策略,即各运营商 P2P cache 系统的目标是最小化本网络为 P2P 流量所支付的费用。在第二种是无私的协作策略,即各运营商 P2P cache 系统的目标是最小化本运营商和对等运营商为 P2P 流量所支付的总费用。虽然上述 2 种策略实现了不同 P2P 缓存系统之间相互协作的目的,但是该策略并不适用于同一运营商的网络内多个域的 P2P 缓存系统之间相互协作的情况。

文献[9,10]提出了一种基于文件全局价值的合作缓存策略,该策略根据文件在各域的流行程度和域间传输文件的链路代价计算每个文件的全局价值(用于衡量缓存或替换文件的收益),当缓存节点需要执行缓存替换时,会优先替换全局价值最小的文件。

2.3 已有缓存策略中存在的不足

当前的协作缓存策略都不考虑缓存节点为本域用户服务和为外域用户服务的差别，只是简单地根据文件在本域用户或全局用户中的流行程度来决定要缓存或替换的文件。但是，P2P 缓存节点为本域用户服务和为外域用户服务所产生的开销是不同的，域间传输数据增大了传输数据的开销，而对用户来讲，域间传输数据也会带来更大的网络延迟。

另外，不同区域的用户的兴趣各不相同，同一文件在不同域的流行程度也存在差别。因此，在设计协作缓存替换算法时，有必要同时考虑文件在不同区域的流行程度和域间传输数据的链路代价对协作缓存系统总收益的影响。

此外，现有协作缓存策略只考虑了获取内容数据的传输开销，而没有考虑存储开销。

3 P2P 缓存系统中开销分析

本节考虑请求文件在各域的流行程度、文件大小、不同域间传输数据的链路开销以及缓存节点缓存数据的存储开销，计算协作缓存系统为全局请求某个文件的所有用户服务所产生的总开销，并在此基础上计算系统缓存该文件带来的缓存增益。

3.1 文件请求开销分析

在协作 P2P 缓存系统中，用户可能从 3 类数据来源获取所请求的文件：本域的缓存节点、邻域的缓存节点或分布在互联网中的 P2P 节点。本文采用单位链路开销来衡量用户从不同数据来源获取文件所产生的传输开销，单位链路传输开销是指在网络中的某条链路上传输 1 byte 的数据所耗费的开销。把在域内链路上传输数据的单位链路开销记为 c_l ，把在 i 域和 j 域之间的链路上传输文件的单位链路开销记为 $c_{i,j}$ ，把在本运营商的网络和其他运营商的网络之间的链路上传输数据的单位链路开销记为 c_e 。通常情况下，用户从其他运营商的网络中获取数据的单位链路开销要大于从本域或邻域的缓存节点获取数据的单位链路开销，即 $c_e > c_{i,j} > c_l$ 。参照文献[9,10]，将这 3 类链路开销进行简化，设 $c_e=1, 0 < c_{i,j} < 1, c_l=0$ 。除了传输开销，还需要考虑缓存节点缓存数据所需的开销，本文用单位缓存开销 c_s 表示单位时间内缓存节点存储单位数据的开销。

根据在不同链路上传输数据的链路开销和缓存节点存储数据的缓存开销，结合用户对文件的访

问情况，可以计算出在一个周期内协作缓存系统为全局请求某个文件的所有用户服务所产生的总开销。它主要和以下因素有关：1) 各域的用户在一个周期内对该文件的请求次数，即该文件的被请求频率，这反映了该文件在各域的流行程度；2) 该文件在全局中各缓存节点的存储情况；3) 不同链路上传输数据的单位链路开销；4) 网络缓存设备存储该文件的缓存开销；5) 该文件的大小。协作缓存系统可以通过调整文件在各域缓存节点的存储情况来使得为全局用户服务所产生的总开销最小。

本文后续讨论中用到的符号对照表如表 1 所示。

表 1 符号对照

符号	描述
A	所有域组成的集合
A^h	本地缓存点已存储文件 h 的域的集合
$A \setminus A^h$	本地缓存节点没有存储文件 h 的域的集合
F_i^h	域 i 的用户在一个周期内对文件 h 的请求次数
c_l	域内传输数据的单位链路开销
$c_{i,j}$	i 域和 j 域之间传输数据的单位链路开销
S_h	文件 h 的大小
c_s	一个周期内存储设备缓存数据的单位存储开销

在协作 P2P 缓存系统中，文件 h 在全局所有缓存节点的存储情况可以分为以下 3 类。

1) 所有域的缓存节点都没有存储文件 h 。在这种情况下，各域中请求文件 h 的用户都需要从互联网中其他 P2P 缓存节点处下载文件 h 。协作缓存系统为一个周期内全局所有请求文件 h 的用户服务所需要的传输开销 C_{delivery} 为 $\sum_{i \in A} F_i^h S_h c_e$ ，存储开销 C_{cache} 为 0。即总开销 C_{total} 为

$$C_{\text{total}} = C_{\text{delivery}} + C_{\text{cache}} = \sum_{i \in A} F_i^h S_h c_e \quad (1)$$

2) 部分域的缓存节点存储了文件 h 。在这种情况下，已缓存文件 h 的域的用户可以直接从本域的缓存节点下载文件 h ，而没有缓存文件 h 的域的用户可以选择从其网络距离最近且已缓存了文件 h 的缓存节点下载文件 h 。协作缓存系统为一个周期内全局所有请求文件 h 的用户服务所需要的传输开销 C_{delivery} 为 $\sum_{i \in A \setminus A^h, j \in A^h} F_i^h S_h \min(c_{i,j})$ ，存储开销 C_{cache} 为 $\text{card}(A^h) c_s S_h$ 。即总开销 C_{total} 为

$$C_{\text{total}} = C_{\text{cache}} + C_{\text{delivery}} = \text{card}(A^h) c_s S_h + \sum_{i \in A \setminus A^h, j \in A^h} F_i^h S_h \min(c_{i,j}) \quad (2)$$

其中, $card(A^h)$ 表示集合 A^h 中元素的个数。

3) 所有域的缓存节点都存储了文件 h , 在这种情况下, 各域的用户都可以从本域的缓存节点下载文件 h , 协作缓存系统为一个周期内全局所有请求文件 h 的用户服务所产生的传输开销 C_{delivery} 为 0, 存储开销 C_{cache} 为 $card(A)c_s S_h$ 。即总开销 C_{total} 为

$$C_{\text{total}} = C_{\text{delivery}} + C_{\text{cache}} = card(A)c_s S_h \quad (3)$$

3.2 缓存对系统开销的影响分析

由 3.1 节分析可知, 在 i 域缓存文件 h 会对系统的总开销产生一定的影响。本文将如果在 i 域的缓存节点存储文件 h 的 1 byte, 那么在单位时间内可以降低的全局总开销称为文件 h 在 i 域的缓存增益 g_i^h 。这里 i 域的缓存节点存储文件 h 后降低的全局总开销是相对于 i 域的缓存节点不存储文件 h 时, 协作缓存系统为全局请求文件 h 的用户服务所产生的总开销来讲的。假设 i 域的缓存节点不存储文件 h 时, 协作缓存系统为全局单位时间内请求文件 h 的用户服务所产生的开销为 C_0 ; 而如果 i 域的缓存节点缓存了文件 h , 协作缓存系统为全局单位时间内请求文件 h 的用户服务所产生的开销为 C_1 , 那么在 i 域缓存文件 h 的缓存增益可以表示为 $(C_0 - C_1)/S_h$ 。结合上面讨论的协作缓存系统为全局请求文件 h 的用户服务所产生的开销, 文件 h 在 i 域的缓存增益可以表示如下。

1) 如果其他域的缓存节点都没有存储文件 h , 那么文件 h 在 i 域的缓存增益为

$$g_i^h = F_i^h c_e + \sum_{j \in A \setminus \{i\}} F_j^h (c_e - c_{i,j}) - c_s S_h \quad (4)$$

式(4)的第一部分表示 i 域的用户可以从本域直接获取文件 h 而不需要去外网下载该文件所节省的开销, 第二部分表示其他域的用户可以从 i 域获取文件 h 而不需要从外网下载该文件所节省的开销, 第三部分表示 i 域存储文件 h 所需要的存储开销。

2) 如果系统中某些域的缓存节点已经存储了文件 h , 那么文件 h 在 i 域的缓存增益为

$$g_i^h = F_i^h \min(c_{i,j}) + \sum_{k \in A \setminus \{i\}} F_k^h [\min(c_{k,j}) - \min(c_{i,j})] - c_s \quad (5)$$

其中, $j \in A^h, k \in \{A^h \cup i\}$ 。

式(5)的第一部分表示 i 域的用户可以从本域直接获取文件 h 而不需要去外域下载文件 h 所节省的开销, 第二部分表示其他域的用户可以就近从 i 域获取文件 h 而不需要再去网络距离更远的其他域获

取文件所节省的开销, 第三部分表示 i 域存储文件 h 所需要的存储开销。

从上面 2 个公式可以看出, 文件 h 在某个域的缓存增益和以下参数有关: 系统中不同域间传输数据的单位链路开销, 文件 h 在各域的缓存节点的存储情况, 各域的用户对文件 h 的请求频率以及文件在 cache 中的单位存储开销。与该域有关的参数(如该域和其他域之间传输数据的单位链路开销, 单位存储开销以及该域的用户对文件 h 的请求频率)的值可以由该域的缓存节点直接测量得到, 而与该域无关的参数的值需要由其他域的缓存节点进行测量。因此, 各域的缓存节点在计算文件的缓存增益时, 需要共享缓存节点所测量的信息。为了方便不同域之间进行信息共享, 可以采用基于 DHT 的全局信息共享系统。尽管各缓存节点之间共享信息会产生一些额外的开销。但是, 由于 P2P 系统中文件的尺寸一般比较大, 相对于传输文件所产生的流量, 缓存节点之间信息交换所产生的开销基本可以忽略。

各域的缓存节点采用下述方式测量与该域有关的参数的值, 并上报到全局信息共享平台, 以便与其他域的缓存节点共享这些信息。

首先讨论如何测量某个域与其他域之间传输数据的单位链路开销。一般情况下, 该参数和网络状况有关, 通常波动不会很大, 缓存节点可以周期性地统计该参数的值。

其次讨论如何测量文件在各域的缓存节点的存储情况。当某个域的缓存节点需要测量本域是否缓存了某个文件时, 只需要查询已缓存文件的记录即可得到该信息。因此, 为了获得文件在各域的缓存节点的存储情况, 每个缓存节点可以维护一份本域已缓存的文件的记录, 同时与其他域的缓存节点共享这些信息即可。

再次讨论如何测量某个域的用户对文件的请求频率。由于缓存节点存储文件的目的是为以后请求该文件的用户服务, 因此这里文件的被请求频率是用来反映文件在未来一段时间的流行程度。然而, 由于用户在未来对文件的请求情况是未知的, 只能通过分析过去一段时间内用户对文件的请求情况来预测。考虑到用户对文件的请求频率在短时间内不会发生较大变化, 所提策略采用本域的用户在最近时间 T 内对文件的平均请求频率作为本域的用户对该文件的请求频率。由于用户对文件的请求

频率会随时间变化, 因此选择一个合适的时间段 T 对预测用户对文件的请求频率至关重要。如果 T 值选取过小, 用户在局部时间段内对文件的请求会对测量结果产生较大的影响, 导致得到的请求频率不准确。如果 T 值选取过大, 测量结果会受较老的用户访问记录影响, 导致所求的请求频率不能实时反映出用户对文件请求情况的变化。

最后讨论缓存节点单位时间内缓存单位数据的存储开销。该参数可以根据缓存设备的相关参数来获得, 如果缓存设备不统一, 可以与其他域的缓存节点共享该参数信息即可。

4 总开销最小的协作缓存策略

本节详细介绍基于总开销最小的协作缓存策略。首先, 介绍基于系统全局总开销的协作缓存系统的工作流程, 该流程是对所提策略的总体描述。然后, 介绍基于系统总开销最小的缓存和替换算法。

4.1 基于文件缓存增益的协作缓存系统工作流程

协作 P2P 缓存系统是由部署在运营商各个网络域中的缓存节点组成。如图 1 所示, 在某运营商的网络中有多个域, 每个域代表一个自治的网络, 可以是一个城域网或省网。各域的网关出口都部署了 P2P 缓存节点, 这些 P2P 缓存节点组成一套全局协作 P2P 缓存系统, 协作存储文件并为来自各域的用户提供服务。

在该结构下, 采用本文所提策略, 各域的缓存节点拦截本域的用户 P2P 请求, 并引导用户从网络距离最近的数据源下载文件, 具体的工作流程如下。当用户的 P2P 请求到达所在域的网关出口, 本域的缓存节点解析该 P2P 请求。如果用户所请求的文件 h 已经被存储在本域的缓存节点, 该缓存节点直接将文件 h 返回给用户。否则, 本域的缓存节点查询文件 h 在其他域的缓存节点的存储情况, 并将用户的请求转发给与用户网络距离最近且已经缓存了文件 h 的缓存节点, 由这个缓存节点将文件返回给用户。如果所有域的缓存节点都没有存储文件 h , 则将用户的请求转发给互联网中的其他 P2P 节点, 由这些 P2P 节点将文件返回给用户。当用户所请求的文件在本域的缓存节点中未命中时, 本域的缓存节点将计算缓存该文件的缓存增益, 并决定是否缓存该文件。基于文件缓存增益的协作缓存流程如图 2 所示。

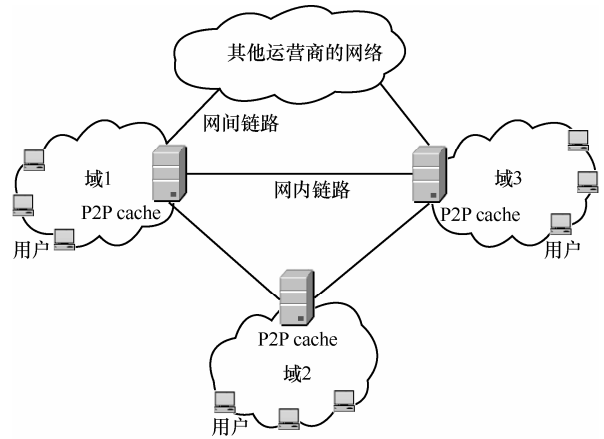


图 1 P2P cache 系统架构

4.2 缓存替换算法

当某个域的用户所请求的文件在本域的缓存节点中未命中时, 该缓存节点将根据文件的缓存增益决定是否存储该文件, 以及替换哪些已经缓存的文件。

在 P2P 缓存系统中, 由于用户是以分片为粒度下载文件, 因此缓存节点在执行缓存替换时, 既可以选择以文件为粒度替换整个文件, 也可以选择以分片为粒度只替换文件的一部分。由于以分片为粒度进行缓存替换可以更充分地利用缓存节点的存储空间, 本文采用以分片为粒度执行缓存替换, 基本思想是用缓存增益较大的分片替换缓存节点中缓存增益较小的分片。

假设 i 域中的用户请求文件 h , 但 i 域的缓存节点没有存储文件 h 。如果本域的缓存节点有足够的剩余存储空间且该文件的缓存增益 g_i^h 大于 0, 则该缓存节点将直接存储文件 h 。否则, 将该文件的缓存增益 g_i^h 与本地已缓存的文件的缓存增益进行比较。如果 g_i^h 是最小的, 本域的缓存节点不存储文件 h 。否则, 本域的缓存节点将用文件 h 替换一个或多个已缓存的缓存增益最小的文件。在执行缓存替换时, 缓存节点首先比较文件 h 和已缓存的缓存增益最小的文件（假设为 h_1 ）的大小, 如果文件 h 的尺寸不大于文件 h_1 , 本域的缓存节点删除文件 h_1 的部分或全部分片, 腾出足够的缓存空间存储文件 h 。否则, 缓存节点比较文件 h 和文件 h_1+h_2 (h_2 为缓存节点中缓存增益第二小的文件) 的大小。如果文件 h 的尺寸不大于文件 h_1+h_2 的尺寸, 本域的缓存节点删除文件 h_1 的所有分片和文件 h_2 的部分或全部分片, 腾出足够的缓存空间存储文件 h 。否则, 本域的缓存节点将继续寻找待替换的文件, 直到所

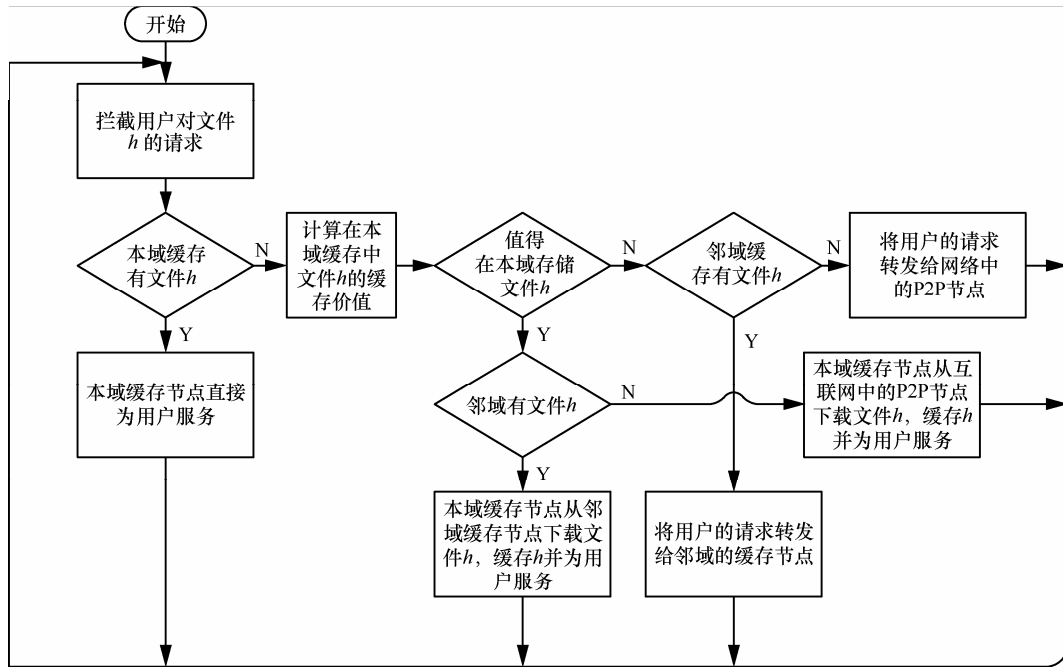


图 2 基于文件缓存增益的协作缓存系统工作流程

有已缓存的文件的缓存增益都比文件 h 的缓存增益大或文件 h 的所有分片都被存储在缓存节点中。由于在 P2P 系统中，用户对文件的请求可以从任意一个分片开始，因此同一个文件的各个分片的缓存增益是相等的，采取随机方式存储或删除同一个文件的分片。所提缓存替换算法的伪代码如图 3 所示。

```

i 域用户请求文件 h，但在本域的缓存节点未命中
1) 计算文件 h 在域 i 的缓存增益  $g_i^h$ 
2) if ( $g_i^h < 0$ ), break;
3) else
4)   if (本域缓存节点有足够剩余空间存储文件 h) then
5)     本域缓存节点直接存储文件 h;
6)   else
7)     if ( $g_i^h <$  本域所有已缓存文件的缓存增益) then
8)       本域缓存节点不会存储 h;
9)     else
10)      对本域已缓存的 N 个缓存增益小于  $V_i^h$  的文件按其缓存增益从小到大排序;
11)      for ( $n=1 \dots N$ ) {
12)        if ( $S_n \leq S_{h_1} + \dots + S_{h_n}$ ) then
13)          用文件 h 替换文件  $h_1$  到  $h_n$ , break;
14)        else
15)           $n++$ ;
16)        end if
17)      end for
18)    end if
19)  end if
20) end if
    
```

图 3 缓存替换算法的伪代码

步骤 1)~3): 计算文件在本域的缓存增益，如果该增益小于零，说明没有缓存价值，不进行缓存

或替换处理，否则进入下一步；

步骤 4)~6): 判断本域缓存节点是否有足够的剩余存储空间，如果有则直接存储该文件，否则进入下一步；

步骤 7)~9): 将该文件的缓存增益与本域所有已缓存文件的缓存增益进行对比，如果该文件的缓存增益小于已有文件的缓存增益，则不进行替换；否则，进入下一步；

步骤 10)~15): 用该文件依次替换缓存增益最小的文件，直到

步骤 16)~20): 算法结束。

由于决定文件缓存增益的几个参数的值会随时间发生变化，在本文所提策略中，各缓存节点将周期地重新计算已缓存的文件的缓存增益。通过这种方式，保证各缓存节点所维护的已缓存的文件的缓存增益可以实时地反映出系统缓存该文件为运营商带来的收益。

5 仿真结果分析

为验证本文所提策略的有效性，通过仿真来衡量协作 P2P cache 系统中所提缓存策略所取得的缓存总收益率。缓存总收益率是指采用缓存系统为运营商节省的开销与不采用缓存系统时用户下载文件所产生的总开销的比率。

基于 OMNET++^[11]，开发并搭建了如图 1 所示

的协作缓存系统。仿真过程中,根据文献[3~10]对协作缓存系统进行如下设置:1)假设P2P网络中共享的文件总数目为5000个,各文件的大小相同;2)假设参与协作的域有8个,且每个域的网关出口都部署了缓存节点,各缓存节点的存储空间容量相同,最多可以缓存所在的域中用户请求的所有文件总大小的20%;3)假设任意2个域之间传输数据的单位链路开销相等且在仿真过程中保持不变;4)假设各域的用户对同一文件的请求频率相同,用户对文件的请求频率反映了该文件的流行程度,二者之间的关系可以用MZip- $f(\alpha, q)$ 分布描述。各域的缓存节点采用本域的用户在最近24h内对某个文件的平均请求频率作为本域的用户对该文件的请求频率,并每隔24h重新统计一次已缓存的文件的缓存增益。

缓存系统所取得的总收益率主要和以下因素有关:参与协作的域的数量,域间传输数据的单位链路开销,以及用户对文件的请求频率的分布。在仿真实验中,分别调整这些参数,来衡量不同的缓存策略所取得的效果。默认情况下,各域的用户对文件的请求频率服从MZip- $f(0.8, 20)$ 分布。

首先,将单位存储开销设为0.005,域间传输数据的单位链路开销从0增大到1,得到如图4所示的系统收益结果。

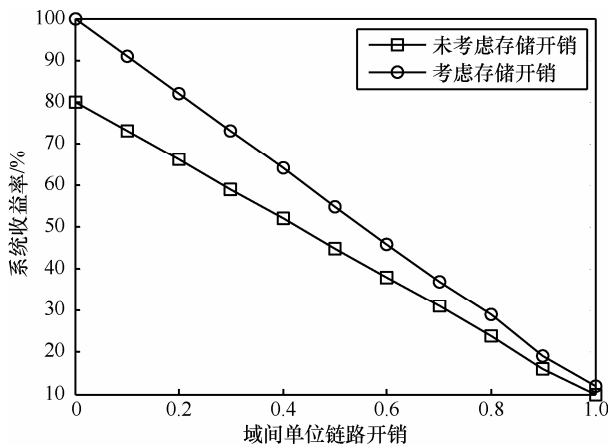


图4 域间单位链路代价变化时所提策略的效果

从图4可以看出,1)不管域间传输数据的单位链路代价如何变化,所提策略所取得的系统收益率都较大。这是由于未考虑存储开销的协作缓存策略只依据传输开销进行缓存和替换,而忽略了存储开销的影响,但是某些情况下(例如文件流行度较低),存储开销可能比传输开销对系统总开销影响

更大。所提策略同时考虑了传输开销和存储开销,系统会根据单位存储开销和域间传输数据的单位链路开销的变化调整文件在缓存系统中的分布,以提高系统收益率。2)随着域间单位链路开销的增大,考虑存储开销前后的收益率差距越来越小。这是因为当域间单位链路开销增大时,传输开销在系统开销中占的比重增加,即存储开销对系统开销的影响降低,因此考虑存储开销前后的系统收益率也随之减小。

然后,将域间传输数据的单位链路开销设为0.5,单位存储开销从0增大到0.01,得到如图5所示的系统总开销结果。

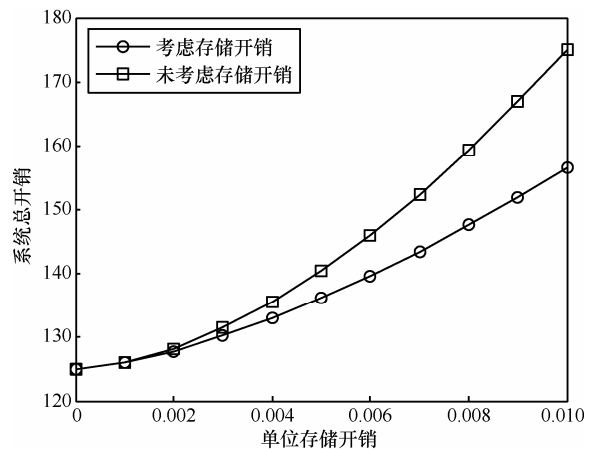


图5 单位存储开销与系统总开销

从图5可以看出,1)考虑存储开销的协作策略比未考虑存储开销的协作缓存具有更小的系统总开销。这是因为未考虑存储开销的协作策略在做缓存决策时只考虑传输开销,只要有空闲缓存空间,系统会对所有的文件加以缓存。但是,在某些情况下(例如文件流行度较低),系统对文件的缓存开销要大于从邻域缓存节点或互联网节点获取文件所需的传输开销,缓存这些文件将会增加系统总开销。而本文所提的考虑存储开销的协作策略通过综合考虑传输开销和缓存开销,将不对这些文件进行缓存,从而具有更小的系统总开销。2)随着单位存储开销的增大,考虑存储开销前后的系统总开销差距越来越大。这是由于不考虑存储开销的协作缓存策略只根据传输开销做缓存决策,而单位存储开销的增大会导致实际系统总开销的增大。本文所提策略则同时考虑传输开销和存储开销,当存储开销增大时,系统会做出不缓存的决策以降低存储开销的影响,降低系统总开销。

6 结束语

本文在深入调研 P2P 缓存技术的基础上,提出了基于全局开销最小的协作缓存策略,该策略综合考虑了数据传输开销和存储开销,并采用文件的缓存增益来衡量缓存节点存储该文件可以为运营商带来的收益。在缓存节点需要进行缓存替换时,优先替换缓存增益最小的文件。仿真结果表明,与目前协作缓存策略相比,该策略可以提高运营商的总收益率和降低系统的总开销。

参考文献:

- [1] YANG H. Design and implementation of P2P flow management system based on cache[A]. 2011 International Conference on Electronic and Mechanical Engineering and Information Technology (EMEIT 2011)[C]. Harbin, China, 2011. 3773-3776.
- [2] 梁卓明, 黄伟强, 郑凯. P2P 流量本地优化综合机制[J]. 计算机系统应用, 2012, 21(1): 94-97.
LIANG Z M, HUANG W Q, ZHENG K. P2P traffic comprehensive mechanism of local optimization[J]. Computer Systems & Applications, 2012, 21(1): 94-97.
- [3] FUJIMOTO T, ENDO R, MATSUMOTO K, *et al.* Video popularity based caching scheme for P2P Video-on-Demand streaming[A]. 2011 IEEE International Conference on Advanced Information Networking and Application(AINA 2011)[C]. Singapore, Singapore, 2011. 748-755.
- [4] MATE M, VIDA R, CSASZAR A, *et al.* Offloading video servers: P2P and/or caches?[A]. The 17th International Telecommunications Network Strategy and Planning Symposium (NETWORKS 2012)[C]. Roma, Italy, 2012. 1-6.
- [5] HEFEEDA M, SALEEH O. Traffic modeling and proportional partial caching for peer-to-peer systems[J]. IEEE Transactions on Networking, 2008, 16(6):1447-1460.
- [6] HEFEEDA M, NOORIZADEH B. On the benefits of cooperative proxy caching for peer-to-peer traffic[J]. IEEE Transactions on Parallel and Distributed Systems, 2010, 21(7): 998-1010.
- [7] DAI J, LI B, LIU F, *et al.* On the efficiency of collaborative caching in ISP-aware P2P networks[A]. 2011 Proceedings IEEE INFOCOM[C]. Shanghai, China, 2011. 1224-1232.
- [8] DAN G. Cache-to-cache: could ISPs cooperate to decrease peer-to-peer content distribution costs?[J]. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(9): 1469-1482.
- [9] WANG B, ZHANG Y, ZHOU X, *et al.* C-LGV: A novel cooperative caching scheme for P2P caches[A]. IEEE International Conference on

Communications(ICC2013)[C]. Budapest, Hungary. 2013. 3012-3017.

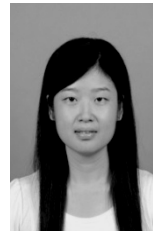
[10] ZHANG Y, ZHOU X, LIU Y, *et al.* A novel cooperative caching algorithm for massive P2P caches[J]. Peer-to-Peer Networking and Applications, 2013, 6(4): 425-433.

[11] OMNET++[EB/OL]. <http://www.omnetpp.org/>, 2013.

作者简介:



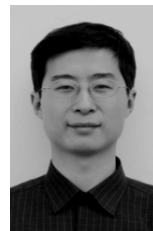
刘银龙(1981-), 男, 安徽宿州人, 中国科学院助理研究员, 主要研究方向为未来网络、移动互联网、无线通信理论。



汪敏(1980-), 女, 湖北孝感人, 博士, 北京城市学院讲师, 主要研究方向为无线通信理论、无线资源管理等。



马伟(1976-), 男, 陕西西安人, 博士, 中国科学院副研究员, 主要研究方向为无线通信理论、移动互联网、大数据分析等。



周旭(1976-), 男, 四川成都人, 博士, 中国科学院副研究员, 主要研究方向为计算机网络、分布式计算。



胡亚辉[通信作者](1982-), 女, 河南信阳人, 博士, 中国科学院副研究员, 主要研究方向为移动互联网、无线传感器网络等。E-mail: huyahui@iie.ac.cn。