

基于网络坐标的P2SP高效下载机制

[刘燕霞 王文琳]

摘要

针对P2SP技术中随机选择策略增大网络流量消耗的问题,提出一种基于网络坐标的机制。在这种机制下,下载节点通过网络坐标的计算选择与自己邻近的服务器进行连接,并且通过服务器得到邻近的节点,使得下载节点能同时从服务器和邻近节点下载资源。仿真结果表明:基于网络坐标的P2SP快速下载机制使得覆盖网络更有效率,降低了开销。



关键词: P2SP 流量 网络坐标 服务器

刘燕霞

女,重庆邮电大学通信与信息工程学院,硕士研究生,研究方向为光通信网络。

王文琳

女,重庆邮电大学通信与信息工程学院硕士研究生,研究方向为光通信网络。

1 引言

近年来,P2SP技术是计算机领域中研究的热点,P2SP下载方式实际上是对P2P技术的进一步延伸,它打破了传统的C/S模式^[1],是一种基于对等节点非中心化服务的平台。P2SP下载方式不但支持P2P技术,同时还通过下载资源数据库把原本孤立的服务器和其镜像资源与P2P资源整合到了一起。这种下载方式克服了用户使用P2S下载在当服务器忙、带宽吃紧或是服务器崩溃的情况下,下载速度慢甚至无法下载的尴尬境遇;同时也克服了用户使用P2P下载时的下载源不稳定和操作上麻烦的缺点,使下载速度更快,下载资源更丰富,下载稳定性更强^[2]。然而P2SP下载机制中下载节点在选择服务器和其他对等节点时是随机的,这种随机选择使得下载节点无法根据自身位置有偏选择服务器和对等节点,使得在为大量用户提供高效而快速的资源链接的同时占用了大量的带宽。

因此,研究如何提高P2SP下载机制的下载性能对降低网络负载有重要的意义。

网络坐标^[3](Network Coordinate)是通过少量的端到端测量来预测网络距离的工具。其核心思想是根据少数几个节点的距离的测量结果,通过特定的优化函数,将该节点映射到特定的几何空间,用节点在几何空间中的坐标表示该节点在网络中位置,通过简单的坐标运算预测任意两个节点间的距离^[4]。这样通过两个节点的网络坐标就可以预测出他们的网络距离。

Vivald^[5]是美国麻省理工学院的F. Dabek于2004年提出的具有代表性的分布式网络坐标系统,已经广泛部署在许多著名的实际系统中,例如SBON^[6]、Bamboo DHT^[7]和Vuze BitTorrent^[8]。为了解决P2SP下载机制中节点随机选择服务器和邻居带来的问题,本文将网络坐标技术应用于P2SP系统,将Vivaldi与P2SP系统进行结合,提

2 NC-P2SP机制的系统设计

以迅雷为代表的下载工具采用的就是P2SP技术，其下载方式的原理如图1所示。P2SP下载的基本过程是：首先客户端(以下称为下载节点)会向服务器发出下载资源请求，之后服务器进行相同资料的检索，并将结果返回给客户端。在这个过程中，服务器将通过多媒体检索数据库把孤立的服务器资源和已存在相同文件的用户(以下称为Peer)即P2P用户资源整合到一起。最后连接服务器所传递的资源地址，所有资源服务器或个人电脑同时向客户端传输下载内容的不同部分，以便加速下载^[9]。

本文设计的NC-P2SP高效下载机制的结构沿用了林源, 陈阳等人的想法^[10], 同样增加了网络坐标系统和开放式数据库系统两个模块。系统的总体结构如图2所示。这两个模块的主要功能分别如图1所示。



通过添加这两个模块，下载节点对服务器的选择不

图2 NC-P2SP系统结构图

Vivaldi把整个互联网看作是一个弹簧系统，网络节点之间的连接相当于弹簧，各个节点通过弹簧相互有力的作用，力的大小和当前预测的距离与实际距离的差值成正比。当网络中各个节点趋向理想位置时，整个系统弹性势能最小，系统趋于稳定。假设 x_i 代表节点 i 的网络坐标，网络坐标的计算就是最小化误差计算公式(1)中 E 的值的结果，通过这个最小化过程来使得弹簧系统能量最小^[5]。

在实际的分布式Vivaldi系统当中，节点 i 有自己的网络坐标 x_i 和本地误差 e_i ，通过多次更新它们不断地接近理想值。在每一次更新中，节点先测量它与某个邻居节点的实际延时，然后通过测得的延时对当前的网络坐标 x_i 和本地误差 e_i 的值进行调整。 c_e 和 c_c 是系统中可调的参数。Vivaldi算法的步骤如下所示：

$$(1) w = e_i / (e_i + e_j),$$

$$(2) e_s = \left\| x_i - x_j \right\| - l_{ij} / l_{ij},$$

$$(3) e_i = e_s \times c_e \times w + e_i \times (1 - c_e \times w),$$

$$(4) \delta = c_c \times w,$$

$$(5) x_i = x_i + \delta \times (rtt - \left\| x_i - x_j \right\|) \times u(x_i - x_j).$$

此算法描述了一个Vivaldi节点的一次坐标更新过程。首先, 根据节点 i 和 j 的 e_i 和 e_j 可以得到一个对于本次更新过程的权重值 w 。然后计算出本地节点和邻居节点之间通过网络坐标预测出来的距离 $\left\| x_i - x_j \right\|$ 和实际测量距离的相对误差 e_s 。利用 w 和 e_s 可以计算出新的本地误差 e_i 。最后, 根据上面的结果, 最终计算出更新后的网络坐标。每个节点加入此系统后, 周期性的运行此算法, 即可逐步缩小预测误差, 达到理想位置。

2.4 NC-P2SP下载系统工作流程

NC-P2SP下载系统的工作说明如下:

(1)在本文的系统中, 所有服务器和Peer节点都加入到分布式网络坐标系统Vivaldi中分别计算自己的网络坐标。

(2)通过开放的分布式数据库OpenDHT^[12], 服务器对网络坐标进行发布。同时下载节点可以事先从OpenDHT获得所有服务器的网络坐标。

(3)下载节点将结合自己的网络坐标和所有服务器和Peer节点的网络坐标, 计算出它到每个服务器和Peer节点的距离, 然后选择最近的服务器进行连接并请求邻居列表。这样每个服务器维护的节点都是与自己相临近的下载节点。

(4)服务器收到连接请求后, 将从自己维护的近邻居列表中随机挑选出 m 个节点, 连同从其他服务器维护的远邻居列表中挑选出来的 k 个邻居节点一同返回给下载节点。最终形成优化后的P2SP下载覆盖网络。

NC-P2SP机制中引入了网络坐标系统, 使得每个下载节点都可以选择离自己最近的服务器进行连接, 然后服务器也能选择离自己较近的节点。这样下载节点不仅能和

最近的服务器进行连接, 同时能通过它获得 m 个同一聚类的近邻居节点。这样不仅节约了它们之间的连接开销, 同时节约了下载节点之间数据传输的开销, 极大地优化了P2SP下载的覆盖网络。

3 仿真实验

3.1 仿真环境

本文通过使用King数据集对系统构建出的覆盖网络的性能进行评价。这个数据集是由互联网上1740个域名服务器两两之间的延迟所组成的矩阵。

利用这个数据集, 通过两个评价标准对系统性能进行评价。在每次试验中, 从网络中随机选择5个作为服务器, 其他1735作为下载节点。在这两种性能评价标准中, 分别独立进行10次实验, 最后取平均值。

3.2 性能评价标准

定义1: 传输开销: 覆盖网络中所有节点间连接的平均延迟。它反映的是在不同网络结构中传输相同数据量消耗网络资源的程度。

定义2: 连接开销: 覆盖网络中所有下载节点与服务器连接的平均延时。它反映了在不同网络结构中连接服务器消耗网络资源的程度。

3.3 仿真结果

分别对远距离邻居数为3、6、9、12、15、18的情况下进行了仿真。同时将其性能和传统的P2SP机制进行了比较。

由图3可以看出NC-P2SP机制的引入减少了节点间的传输延时。这说明传输相同的数据量, 采用NC-P2SP机制消耗更少的网络资源。

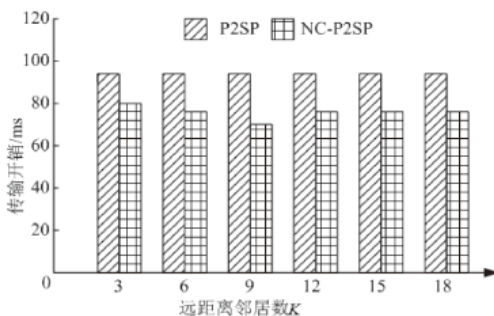


图3 传输开销对比

图4主要对节点连接服务器的平均延时进行了比较。NC-P2SP机制减少了约40%的开销,大大降低了用于连接服务器的网络资源消耗。

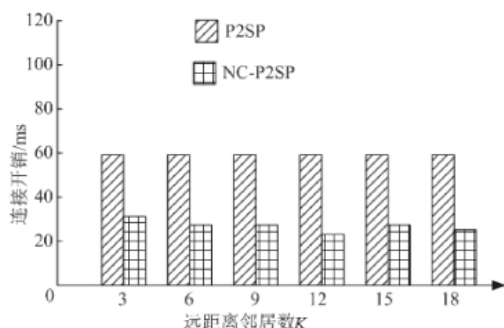


图4 连接开销对比

4 结论

针对P2SP系统的不足,本文提出了基于网络坐标的P2SP下载机制。通过仿真,NC-P2SP下载机制能利用节点本身的计算能力,通过计算网络坐标实现服务器和Peer的有偏选择,优化了覆盖网络,同时减少了下载中用于节点和服务间通信以及节点之间数据传输的开销。这种改进不仅提高了下载服务性能,同时降低了网络流量负载,具有很好的实际意义。

参考文献

- 1 A Parker. The True Picture of Peer-to-Peer File Sharing[EB-OL]. 2007-08-01. <http://www.cachelogic.com>.
- 2 许东梅. 基于CDN的P2SP下载系统的研究与实现[D]: [硕士论文]. 南昌: 江西师范大学计算机信息工程学院 2008.

- 3 Pietzuch P, Ledlie J, Mitzenmacher M, et al. Network-Aware Overlays with Network Coordinates [C] // International Workshop on Dynamic Distributed Systems, Lisbon, Portugal, 2006: 12-12.
- 4 赵小茜. 非中心式网络坐标系统安全问题的研究[D]: [硕士论文]. 北京: 清华大学电子工程系 2009.
- 5 Dabek F, Cox R, Kaashoek F, et al. Vivaldi: A decentralized network coordinate system [J]. ACM SIGCOMM Computer Communication Review, 2004, 34(4): 15 - 26.
- 6 Pietzuch P, Ledlie J, Shneidman J, et al. Network-Aware Operator Placement for Stream-Processing Systems. Proceedings of the 22nd International Conference on Data Engineering (ICDE' 06), Washington, DC, USA: IEEE Computer Society, 2006.
- 7 Rhea S, Geels D, Roscoe T, et al. Handling churns in a DHT. Proceedings of the annual conference on USENIX Annual Technical Conference (ATEC' 04 :), Berkeley, CA, USA: USENIX Association, 2004.
- 8 Azureus-now called Vuze-Bittorrent Client, 2009. <http://azureus.sourceforge.net/>.
- 9 周迪民. 基于ACL控制校园网P2SP下载流量的研究与实现[J]. 计算机系统应用. 2010, Vol.19, No.10.
- 10 林源, 陈阳, 张增斌, 邓北星, 李星. 基于网络坐标的Multi-Tracker机制研究[J]. 清华大学学报. 2009, Vol.49, No.10.
- 11 Sean Rhea, Brighten Godfrey, Brad Karp, et al. OpenDHT: A Public DHT Service and Its Uses [J]. ACM SIGCOMM Computer Communication Review, 2005, 35(4): 73 - 84.
- 12 Gummedi K P, Saroiu S, Gribble S D. King: estimating latency between arbitrary internet end hosts [J]. A CM SIGCOMM Computer Communication Review, 2002, 32(3): 11 - 11.

(收稿日期: 2011-09-21)

征稿启事

本刊长期面向国内外通信行业管理人员、通信企业管理人员、信息通信网工程技术人员、通信市场营销人员以及其它通信产业跟踪研究人员征稿。征稿范围如下:

- ① 通信产业宏观政策分析、通信行业管理信息综合报道、通信产业发展环境研究与分析;
 - ② 信息通信网及相关技术热点跟踪与报道、信息通信网络规划设计、信息通信技术研究;
 - ③ 通信市场跟踪研究与报道、通信企业管理模式及方法探讨、通信企业运行机制研究;
 - ④ 其它与通信行业相关的跟踪、研究与分析报道。
- 来稿一经确定采用即通知作者本人,请投稿者留下准确的联

系方式(联系电话, E-mail, 通信地址等)。编辑部收到投稿即进入稿件处理程序,如欲转投其它刊物,作者应事先与编辑部联系并取得同意,投稿3个月后如未收到刊用通知,可自行处理。限于编辑部人力,来稿一律不退,不发退稿通知,请作者自留底稿。

文责自负,作者应保证无署名争议,无一稿多投,无泄密之处,无政治性问题。严禁抄袭剽窃,文稿中摘编或引用他人作品,请在参考文献中列出其作者和文献来源。

非编辑部约稿,一经采用,将按照国家相关规定收取版面费,具体事宜与编辑部联系。

来稿请发电子邮件至 gdtx@gsta.com。

编辑部来稿处理联系电话: (020) 38639258 方小姐。