

# 一种自适应的 P2P 流媒体数据调度算法

孙名松<sup>1</sup>, 周红敏<sup>2</sup>, 唐 亮<sup>2</sup>

(1. 哈尔滨理工大学 网络信息中心, 哈尔滨 150080; 2. 哈尔滨理工大学 计算机科学与技术学院, 哈尔滨 150080)

(sxlzfm@163.com)

**摘 要:**针对现有算法在有限带宽情况下数据块易丢失的问题,提出了一种新颖的数据调度算法——最少优先调度算法(LFSA)。该算法易于实现,并且可以根据网络的异构性,带宽抖动以及 P2P 网络节点的动态性进行自动调整。实验结果表明,LFSA 能够有效缩短系统播放启动延迟,而且可以在不知道伙伴节点带宽和没有任何负载平衡机制的情况下取得比传统策略更好的性能。

**关键词:**P2P 网络;流媒体;数据调度;启动延迟

**中图分类号:**TP393 **文献标志码:**A

## Adaptive data scheduling algorithm for P2P media streaming

SUN Ming-song<sup>1</sup>, ZHOU Hong-min<sup>2</sup>, TANG Liang<sup>2</sup>

(1. Center of Network Information, Harbin University of Science and Technology, Harbin Helongjiang 150080, China;

2. College of Computer Science and Technology, Harbin University of Science and Technology, Harbin Helongjiang 150080, China)

**Abstract:** In order to solve the problem of losing data block under the limited bandwidth in existed algorithms, a novel scheduling algorithm was proposed for P2P network, named Least-First Scheduling Algorithm (LFSA), which was simple to implement. Furthermore, it can adjust itself automatically to adapt to network heterogeneity, bandwidth fluctuation, and overlay dynamics. Experimental results show that LFSA can shorten the start delay time and achieve better performance than conventional strategies without any knowledge of the bandwidth of its partners or any special mechanism for the load balance.

**Key words:** P2P network; media streaming; data scheduling; startup latency

## 0 引言

近年来,基于 P2P 技术的流媒体应用已成为研究热点。早期的 P2P 方案如 ESM(End System Multicast)<sup>[1]</sup>, NICE<sup>[2]</sup>, ZIGZAG<sup>[3]</sup>等,采用树型结构不能很好地适应节点的动态性和网络的异构性。PRO(Peer to peer Receiver-driven Overlay)<sup>[4]</sup>, DONet(Data-driven Overlay Network)<sup>[5]</sup>等通过 Gossip 协议构造无结构的覆盖网络,提高了系统的稳定性和鲁棒性,并且一些基于 DONet 的实际 P2P 流媒体系统如 PPLive, PPStream 和 Gridmedia,使用一个带宽仅为 5~10 Mbps 的流媒体服务器,就能支持千万级用户同时观看一个热点直播节目。

在基于 DONet 的系统中,媒体流被分为一系列小的数据块,并且在不同邻居节点之间进行交换,系统性能主要依靠覆盖网的构造和数据调度算法。研究者们已经提出许多基于 Gossip 的算法构造覆盖网以提高节点的随机性和可靠性<sup>[6,7]</sup>,但很少有人重视数据调度算法的研究,这对系统性能的影响也是非常重要的。本文通过对现有调度算法存在的问题进行分析,提出了一种新颖的数据调度算法,并且在 Polycast 系统中实现。

## 1 相关研究

目前已经有一些对 P2P 数据调度策略的相关研究。如随机策略(Random)<sup>[8]</sup>,最少优先策略(Rarest-first)<sup>[5]</sup>和循环鲁棒策略(Round-robin)<sup>[9]</sup>等已经被提出且在许多流媒体系

统中采用。

Random 是一个非常简单的 P2P 数据调度策略,它被 Chiansaw<sup>[8]</sup>采用。根据这种策略,对于每个空缺的数据块,每个节点随机地从持有此数据块的伙伴节点中选择一个节点,然后请求选中的伙伴节点。这个策略可以在基于 DONet 的系统中使用,但是其性能不稳定,尤其是在异构的网络环境中。

RR 策略被用在分层流媒体系统中,如 PALS<sup>[9]</sup>。根据这种策略,按照比例分配所有请求的数据块到一个伙伴。如果仅有一个伙伴持有该数据块,那么就从该伙伴获取;否则,从拥有最大可用带宽的伙伴节点获取。这个策略可以取得很好的负载平衡。

RF 是一种快速时间响应的启发式算法,它是在 CoolStreaming/DONet<sup>[5]</sup>中被提出。这种策略的主要思想包括:首先请求潜在提供者较少的数据块;如果一个数据块有多个潜在提供者,那么具有最多剩余带宽和足够可用时间的伙伴将被选中。然而,这种策略有一个内在的缺点,我们将在下一节详细分析。

## 2 自适应的数据调度算法

### 2.1 存在的问题

在 CoolStreaming/DONet 中,调度策略被归结为并行机调度问题(Parallel Machine Scheduling)的一个变种,属于 NP 类问题。因此,在流媒体系统当中,可以取得近似的最佳调度。但是,它在实时性要求较高的流媒体应用中执行得不是很好。在 Polycast 中,一个数据块有很少的潜在提供者,其原因有:

收稿日期:2007-10-10;修回日期:2007-12-07。

作者简介:孙名松(1963-),男,黑龙江哈尔滨人,教授,博士,主要研究方向:网络安全、网络应用;周红敏(1980-),男,山西临汾人,硕士研究生,主要研究方向:网络安全、P2P 网络;唐亮(1980-),男,江苏连云港人,硕士研究生,主要研究方向:网络安全、P2P 网络。

1) 该数据块有一个较小的序号,并且它在一些伙伴中已经滑出了缓冲区滑动窗口;2) 它是一个非常新的数据块,有一个很大的序号,并且没有被广泛传播,因此仅有数据源节点和与数据源直接相连的伙伴会拥有这个数据块。对于第一种情况,使用 RF 算法首先请求有较少潜在提供者的数据块是适合的。但是,对于第二种情况,在带宽不足的情况下,首先请求有较少潜在提供者的数据块可能会导致带有较小序号的数据块将在播放截止时间以前不能到达。除此之外,RF 策略需要提前知道每个伙伴节点的精确带宽,而这在实时环境中是很难测量的。

## 2.2 算法描述

在 Polycast 系统中,把流媒体数据分割成长度相同的多个 TF (Transfer frame) 帧,然后放入发送源的缓冲区里面,每帧长度为 TF\_len 字节,按 TF\_sequence\_number 在缓冲区中顺序存储。在节点缓冲区中 TF 帧的可用性用 Buffer Map (BM) 来记录。BM 另外还记录着当前节点中所存在的最高序列号 maxseq 和最低序列号 minseq 的 TF 帧。并且设定 BM 长度为  $\text{maxseq} - \text{minseq} + 1$  Bits, 每一个 bit 对应表示 TF 帧的可用性,若可用设为 1, 否则设为 0。

为了获得最小化启动延迟和更为流畅的播放体验,调度算法应该首先获取带有较小 TF\_sequence\_number 的 TF 帧。同时,算法必须能够适应高度动态的网络条件并且能取得负载均衡。为了满足这些要求,提出了一种自适应的数据调度算法:最少优先调度算法 (Least-First Scheduling Algorithm, LFSA)。

在 LFSA 中,假定带有较小 TF\_sequence\_number 的 TF 帧具有较高的优先级。

具体算法描述如下:

- 1) 如果一个 TF 帧有较小的 TF\_sequence\_number 且仅有一个提供者,则请求它;
- 2) 否则,它将被插入 buffered\_data 集合;
- 3) 在调度的每个循环中,首先考虑 buffered\_data 中的 TF 帧然后才考虑其他的;
- 4) 如果一个 TF 帧对应多个潜在的提供者,则带有最小 req\_num 的提供者将被选中;
- 5) 如果仍有两个或多个提供者带有相同的 req\_num,那么和当前节点 IP 地址最匹配的节点将被选中。

在以上描述中,req\_num 是预先定义好的一个变量,并初始化为 0。对于每个节点而言,当有其他伙伴节点从它请求一个 TF 帧时,req\_num 加 1; 当它成功发送出去一个 TF 帧时,req\_num 减 1。通过这种策略,调度算法将基于流媒体数据的实际传输状态来调整,并且根据 LFSA, 每个节点总是能快速地有能力提供数据的伙伴获取数据。换句话说,有足够资源(如网络带宽)的伙伴节点总是能提供更多的数据给其他节点。算法的计算复杂度为  $O(M \times N)$ , 这里  $M$  是缓冲区的滑动窗口中 TF\_sequence\_number 的最大值,  $N$  是这个节点的伙伴节点数目。

一个伙伴节点可能在任何时刻只接收而不提供数据,并且它的访问带宽可能随时间而变化,因此需要一个重试机制来重新请求已经请求但很长时间还没接收到的 TF 帧。然而,太多的重新请求可能会引起网络拥塞并导致更多的 TF 帧在播放截止时间前不可达。为了避免这种现象发生,我们定义了一个变量 max\_req, 它是指这个节点可以同时从一个伙伴请求的 TF 帧的最大数,这个变量可以根据 TF 帧的实际大小进行调整,如果 TF 帧尺寸过大,应该减小 max\_req 以避

免由于冗余的重新请求引起的网络拥塞; 否则, 如果 TF 帧尺寸过小, 可以增大 max\_req 以充分利用网络带宽。

## 2.3 实例分析

按照以上描述,在每个节点中,LFSA 根据实际数据传输状态选择合适的数据提供者,因此在异构网络中特别有效。图 1 是数据调度的一个例子。

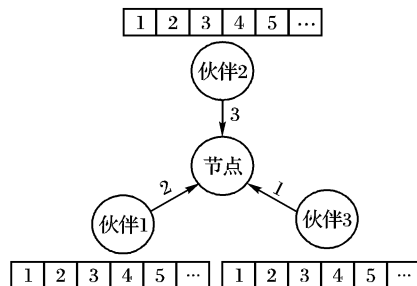


图1 一个数据调度例子

如图 1 所示,中间节点有三个伙伴节点,箭头上的数字代表从伙伴节点到此节点的可用带宽。每个伙伴节点旁边的矩形中的数字代表它持有的 TF 帧情况。伙伴 2 的 IP 地址和中心节点的最为匹配,其次是伙伴 1,最后是伙伴 3。

为了更好地说明问题,在这个例子中我们设置 max\_req 为 2。根据 LFSA 策略,中间节点同时从每个伙伴节点最多请求 max\_req 个 TF 帧,此时无论这个伙伴是否已经成功发送出预先请求的 TF 帧,此节点都会从其他伙伴重新请求下一个 TF 帧。剩余带宽较多的伙伴节点发送一个 TF 帧花费的时间较少,因此将从这些伙伴节点请求更多的 TF 帧。在这个期间,此节点分别从伙伴 1,2,3 请求了 6,9,3 个 TF 帧,并且这个比例正好符合可用带宽,这也说明 LFSA 算法是一种自适应的调度策略。

## 3 性能评价

为了更好地分析 LFSA 算法的性能,在 Polycast 系统中也实现了 RF 调度策略以进行公正地比较。前面提到,我们的目标是最小化播放启动延迟,因此将缓冲相同数量的 TF 帧所用的时间作为评判每种算法性能的主要指标。并且,每个实验结果取 10 次相同实验的平均值。在实验中,覆盖网中都有 50 个节点,每个节点有 4~6 个伙伴。调度算法周期性地执行,我们设定这个周期为 1 s。实验中使用的媒体流编码率为 500 Kbps,每个 TF 帧的尺寸大约为 100 K, max\_req 的值设为 5。

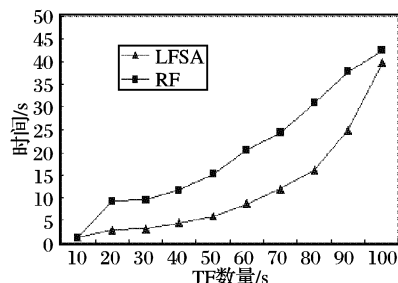


图2 调度策略性能比较

图 2 给出了 LFSA 和 RF 策略的曲线。从图 2 可以看出,为缓冲 60 s 的播放数据,LFSA 仅花费 10 s,而 RF 花费 20 s,而且在大多数情况下,缓冲相同播放时间的数据,LFSA 需要的时间都比 RF 少。由此可见,LFSA 算法明显优于 RF 算法策略。RF 的曲线在开始时有个波动,这是因为它请求了一些新的数据块。

以上实验结果表明,LFSA 可以缩短播放启动延迟,并且

(下转第 567 页)

线 Ad Hoc 网络通信中是不存在的,所以在无线局域网中得到的吞吐率要稍微小一些。当目的节点不在发送节点的传输范围内时,接入点需要向目的节点转发数据帧,但是不需要再次竞争信道。从表 2 可以看出,这时 ExDCF 的吞吐率还是要比标准 DCF 稍高。

当无线网络中有多于一对节点时,ExDCF 的性能提高会变得非常明显。网络中存在多对节点的复杂场景的仿真结果见表 3。

表 3 复杂场景的仿真结果 kbps

场景	移动站点数量		
	5 pairs	10 pairs	15 pairs
Adhoc_DCF	180.3	90.4	60.2
WLANs_DCF	35.3	7.2	3.5
WLAN_ExDCF_near	176.3	89.3	59.3
WLAN_ExDCF_far	96.7	47.6	37.0

由表 3 可见,ExDCF 可以有效提高网络的实际吞吐率,并且随着无线节点数量的增加,这种性能改善越明显。由表 2、3,当网络中只有一对无线节点时,WLAN\_ExDCF\_far 场景下的吞吐率比 WLAN\_DCF 场景下的吞吐率高约 15%,而当网络中有 5 对无线节点时,WLAN\_ExDCF\_far 场景比 WLAN\_DCF 场景下的吞吐率高约 170%。造成这种情况的原因是,随着网络中节点增多,接入点可用带宽变少,同时发生碰撞和重传的概率增大;而在 ExDCF 协议中,接入点转发数据帧时不需要再次竞争带宽,因而接入点不再是网络瓶颈,从而大大改善了网络性能。

## 5 结语

根据无线局域网自身的特点(所有的无线节点都在 AP 的传输范围内),提出了一种对 DCF 协议的改进算法 ExDCF。与 DCF 相比,ExDCF 能有效减少接入点转发数据帧的时间花费,提高带宽利用率。从理论分析和仿真结果都可以看出,ExDCF 在无线局域网中可以提高无线信道的利用率和无线网络的整体性能,而且使无线节点间具有更好的公平性。

(上接第 559 页)

使得播放器在有限制情况下可以更为流畅,特别是在每个伙伴节点带宽不足的情况下,但它们的可用带宽对播放全质量的媒体流是足够的。LFSA 可以保证更好的 QoS,同时,它是完全分布式的,并且可以在不知道伙伴节点带宽和没有任何负载均衡机制的情况下取得比传统策略更好的性能。

## 4 结语

通过和传统的 P2P 流媒体调度算法进行比较,分析了目前流行的调度算法存在的问题,提出了一种自适应的 P2P 流媒体数据调度算法 LFSA。该算法既克服了传统策略的缺点,又能够根据网络的异构性和节点的动态性自动调整调度策略。实验结果表明,LFSA 算法不仅能够有效缩短系统的启动延迟,而且能够改善终端用户的视频质量。

### 参考文献:

- [1] CHU Y H, RAO S G, SESHAN S, *et al.* A case for end system multicast [J]. ACM SIGMETRICS Performance Evaluation Review, 2000, 28(1): 1-12.
- [2] BANERJEE S, BHATTACHARJEE B, KOMMAREDDY C. Scalable application layer multicast [C]// Proceedings of ACM SIGCOMM. Los Angeles: ACM Press, 2002: 43-51.

### 参考文献:

- [1] ISO/IEC 8802-11: 1999(E). IEEE standard for wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications [S]. 2001.
- [2] RFC 2205, Resource ReSerVation Protocol (RSVP) — Version 1 Functional Specification [S]. IETF, 1997.
- [3] LEE S-B, AHN G-S, ZHANG X, *et al.* INSIGNIA: an IP-based quality of service framework for mobile Ad Hoc networks [J]. Journal of Parallel and Distributed Computing, 2000, 60(4): 374-406.
- [4] LEE S-B, AHN G-S, ZHANG X, *et al.* INSIGNIA, draft-ietf-manet-insignia-01.txt [S]. Work in Progress, 1999.
- [5] TERZIS A, SRIVASTAVA M, ZHANG L. A simple QoS signaling protocol for mobile hosts in the integrated services Internet [C]// Proceedings of IEEE INFOCOM'99. NY: IEEE Press, 1999: 1011-1018.
- [6] TSENG C C, LEE G, LIU R-S. HMRSVP: a hierarchical mobile RSVP protocol [C]// Proceedings of International Conference on Distributed Computing Systems Workshop 2001. Arizona: Wireless Networks, 2001: 467-472.
- [7] BIANCHI G. Performance analysis of the IEEE 802.11 distributed coordination function [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(3): 535-547.
- [8] BONONI L, CONTI M, GREGORI E. Design and performance evaluation of an asymptotically optimal backoff algorithm for IEEE 802.11 Wireless LANs [C]// Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 8. Washington, DC: IEEE Computer Society, 2000: 8025.
- [9] WANG G-H, SHU Y-T, FAN Y-J. An efficient resource reservation mechanism in Ad Hoc networks [C]// Canadian Conference on Electrical and Computer Engineering. Montreal: IEEE Press, 2003.
- [10] ZHANG LIANG, SHU YAN-TAI, LIU YAN, *et al.* Adaptive tuning of distributed coordination function (DCF) in the IEEE 802.11 to achieve efficient channel utilization [C]// Future Telecommunications Conference 2003, Beijing, China, 2003.
- [11] Ns-2 network simulator [EB/OL]. [2007-09-01]. <http://www.isi.edu/nsnam/ns/>.

- [3] TRAN D A, HUA K A, DO T T. Zigzag: an efficient peer-to-peer scheme for media streaming [C]// Proceedings of IEEE INFOCOM 2003. San Francisco: IEEE Press, 2003: 1283-1292.
- [4] REJAIE R, STAFFORD S. A framework for architecting peer-to-peer receiver-driven overlays [C]// Proceedings of the ACM NOSS-DAV. Kinsale: ACM Press, 2004: 42-47.
- [5] ZHANG X Y, LIU J, LI B, *et al.* Coolstreaming/DONet: A data-driven overlay network for peer-to-peer live media streaming [C]// Proceedings of IEEE INFOCOM. Miami: IEEE Press, 2005: 2012-2111.
- [6] LIANG JIN, NAHRSTEDT K. Randpeer: Membership management for QoS sensitive peer-to-peer applications [C]// Proceedings of IEEE INFOCOM. Barcelona: IEEE Press, 2006: 1-10.
- [7] VISHNUMURTHY V, FRANCIS P. On heterogeneous overlay construction and random node selection in unstructured P2P networks [C]// Proceedings of IEEE INFOCOM. Barcelona: IEEE Press, 2006: 1-12.
- [8] PAI V, KUMAR K. Chainsaw: Eliminating trees from overlay multicast [C/OL]. [2007-09-01]. <http://mnl.cs.sunysb.edu/home/vinay/papers/chainsaw-iptps.pdf>.
- [9] AGARWAL V, REJAIE R. Adaptive multi-source streaming in heterogeneous peer-to-peer networks [C]// PIE/ACM MMCN'05. ACM Press: [s. n.], 2005: 102-109.