

一种 PUSH 技术支持的分布信息共享机制

A Distributed Information Sharing Mechanism Supported By Another Push Technology

何 军 周明天

(电子科技大学计算机科学与工程学院 成都 610054)

Abstract This paper presents a distributed information sharing mechanism supported by another push technology and describes its principle, feature and key technology. This mechanism uses CATV high speed broadcasting downward channel to push the network information into personal computer, so that changes query method from remote interaction to close interaction while prevents channel congestion from a huge numbers of user accessing network.

Keywords PUSH, CATV, Information sharing, Content query

Internet 建立于 TCP/IP 基础之上,具有良好的开放性和互操作性,它将世界各地的异构网络平滑地连接起来,使用户间无论在何时、何地都可以方便地相互交流,互相协作。

WWW 通过 Homepage、HTML、CGI、Browser 等技术实现了 Internet 上异构信息系统之间的无缝联接,使用户可以方便地共享信息并进行互操作。

Internet 和 WWW 良好的开放性和互操作性使它们受到广泛欢迎,其规模不断迅速扩大,用户数已达一亿,主机5000多万台,网上信息量呈爆炸性增长^[1]。

WWW 上信息查询主要是依据 URL (Uniform Resource Locator) 原理,它提供搜索引擎为用户找来 URL 地址,再由用户根据得到的信息制定查询策略、选择查询路径,最后建立链接进行查阅,从用户开始查询任务直到得到结果的时间等于搜索时间+查询时间+传输时延,随着网络规模的不断扩大,网络阻塞现象日益严重,信息量大和通信信道容量小的矛盾日益尖锐,上述三种时延也相应增加,甚至使用户难以忍受而撤销查询任务。

因此,有必要建立一种新的信息共享机制,它应该具有:与网络规模、用户数几乎无关,响应时间短,按内容查询的优点,这就是采用另一种 PUSH 技术的分布信息系统——我们称之为广存系统。

1 PUSH 技术

PUSH 技术即信息推送技术,它按照一定的技术标准或协议,将网上的信息源上的信息及时、主动地通过宽带广播信道发送给用户,使用户不必上网便可查询到信息。PUSH 技术现已成为第三代浏览器的核心技术,是当前网络竞争的焦点,其标准尚在形成中。

采用 PUSH 技术,使被动、按要求查询的服务器转变为主动的多播或广播服务器。用户可以把更多的精力放在他所关心的内容上,而不必按 URL 地址去大海捞针。信息浏览方式由单一在线、实时交互方式扩展到与离线、非对称交互和单向广播接收方式并存、互补,极大地降低了费用、减少了时延、缓解了信道拥挤。

这种信息获取方式的改变,使得网上信息内容的组织、检查、管理有了技术支持,可以阻断信息垃圾的传播。另一方面,拥有网络信息推送的控制权,也就拥有了控制传播渠道、选择传播内容、决定传播接受者的能力。这既体现了 PUSH 技术的重大意义,也说明了其潜在的风险性。

2 分布式多媒体信息广存系统

本文提出了一种用于分布式多媒体信息广存系

何军 博士生,主要研究方向:计算机网络及应用技术,周明天

教授、博士生导师,院长,主要研究方向:计算机网络,分布对

象技术,并行分布处理和系统集成。

统^[3]的信息共享机制 DMIBSS (Distributed Multimedia Information Broadcasting Storage System), 如图1所示。

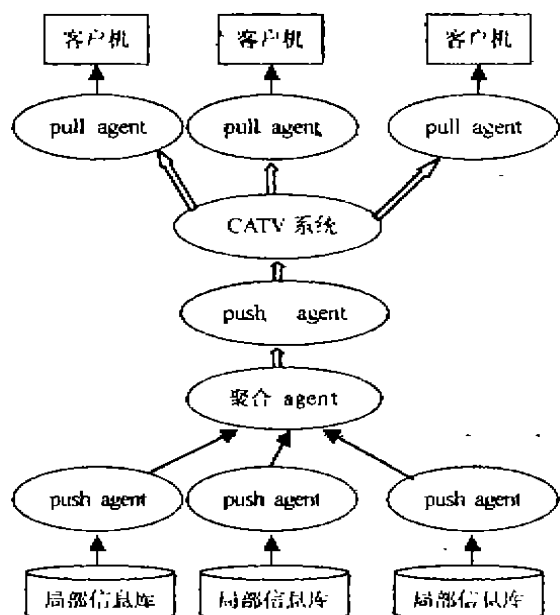


图1 分布式多媒体信息广存系统

局部信息库中可以存放多种类型的信息,各局部信息库可以是异构的,它们通过本地 pull agent 将要发布的信息以一种统一的标准推送给聚合 agent。聚合 agent 将各节点推送来的信息过滤、分类、重组后,并由 push agent 将信息推送给 CATV 系统。CATV 系统经其高速宽带下行有线电视电缆将信息广播推送给用户的客户机。客户机上的本地 push agent 根据用户的需求、兴趣范围等知识,将海量信息拉入客户机的 GB 硬盘中,供用户随时查询。DMIBSS 具有如下一些特点:

- 通讯方式采用高速下行广播信道和低速上行信道的非对称双工模式。

- 采用无连接服务方式。信源只是将信息通过广播信道进行发送,并不指定明确的接收者,用户根据其需要决定是否从信道中抓取 (pull) 数据。因此通讯中不必进行流控、确认重传,当出现数据错误时,由信源的信息循环重发机制来纠正。

- 由于传播方式的对象是面而非点,所以广存系统内信源实体与信宿实体间的工作模式是点一面模式。

- 客户机通常都在忙/闲状态之间转换,因此广存机制就可利用客户机处于空闲状态时将信息推入

客户机,实现悠闲传送服务。

3 实现 DMIBSS 的关键技术

由于 DMIBSS 提供了一种新的信息共享方式,即尽量将远交互方式改为近交互方式,因此对实现它的软硬件技术有新的要求,主要有:GB 硬盘、CATV 系统、服务质量、内容安全检查、基于内容的信息管理技术等。

3.1 GB 硬盘

GB 硬盘是 DMIBSS 的基本硬件配置之一,它用来存贮从 CATV 信道中抓取进来的大量信息供用户使用。硬件技术的发展使 GB 硬盘的容量不断上升,而价格却不断下降,性能价格比不断提高,可以为广大普通用户所接受。

3.2 CATV 高速下行信道

依据 64-QAM (正交调幅技术) 调制的 DVB-C 标准,一个带宽为 8MHz 的 PAL 体制的电视频道,当符号频率为 6MHz 时,可获得 36Mbps (或 16.2 GB/小时) 的下行传输速度。即使扣除 25% 的信道传输开销,净推送能力也超过 12GB/小时。用它广播 1GB 的海量内容只需 5 分钟。一个电视频道可以分成若干个子信道,每个子信道出租给信息发布者自主经营。CATV 广播信道是一种可供用户无冲突共享的信道,用户可以从信道中并行获取数据,不会造成信道拥塞,因此广播结构很适合于用户规模持续扩张的情况。

3.3 服务质量 (QoS)

由于通过 CATV 信道传输的多媒体数据有文本、图形、图像、声音、视频等多种形式,它们各有自己的传输质量要求。同时在传播过程中,由于各种自然物理原因 (如信道信噪比、电缆接触、发送速率、温度和天气等) 会造成数据误码,因此在这种异步广播模式下,为进行数据纠错就需要有新的 QoS 机制。

对待播数据确定优先级,可以保证重要信息优先广播,对要求高速发送的数据可将其放在若干个子信道上并行发送。这些均由 push agent 与 CATV 系统协商解决。通常数据包是采用发确认帧、出错重传的方式来保证正确的,而广存系统是采用循环发送方式来纠错。即当信宿端发现数据出错时,只是丢弃错误数据,等待下次发送时间的到来。另外,由于广存系统发送的是海量数据,而用户的时间总是有限的,不可能全部阅读,因此可以对错误数据简单丢弃忽略不计。随着信道质量的提高,误码率将愈来愈小,相对于海量视像数据,所丢弃的一小部分数据不

会影响用户的阅读效果。

3.4 信息安全

Internet 是无政府的社会,每个人都可以在上面享受到高度的自由,这就使得网上信息良莠不齐,信息垃圾充斥于网上并可不受限制地到处传播。

广存系统是通过多级接力式的 push agent 发布信息的,这就使得信息的发布可以得到控制,过滤掉信息垃圾。每一级 push agent 都设有一个可信度,用以供下一级 push agent 进行信息过滤。当某 push agent 的可信度高时,由它推送的信息可直接向前推送;反之,则需对信息进行按内容检查,符合要求的才能向前传送。

3.5 基于内容的超媒体信息管理

超媒体使得对同样的内容可以用文本、图像、声音、视频等多种媒体形式表达,它们都属于复杂数据,因此把它们以内容为中心进行链接、存贮、查询及其它管理是一种有效的方法。信源将要播送的内容以一种或几种媒体形式表达出来,并按内容建立基本链。客户机按用户的需求从 CATV 信道中自动抓取数据,进行裁剪,存入自己的 GB 硬盘中。硬盘中存贮的媒体是物理媒体,相应地,在用户查询时按内容生成的临时的媒体形式为逻辑媒体。这样用户查询时,可在基本链的基础上按内容生成动态链,并

可以根据用户的要求用不同的逻辑媒体形式作为查询结果。这样,就使以内容为中心的超媒体信息管理具有极大的灵活性、适应性和动态性。

结束语 本文提出了一种新的信息共享方式——PUSH 技术支持的分布信息共享机制。该模型有效地利用了计算机网 IP/PUSH 技术、电视网广播信道和高带宽的优点。在用户提出查询要求前,数据已被预读入 GB 硬盘中,使用户从提出要求到得到结果的响应时间大大缩短。目前,我们已经突破了 10×10^6 bps 在普通 CATV 信道中传输及家庭数据库软件等关键技术,通过了 60 户网软、硬件的实际测试,下一步将进行 300 户网及更高速率数据传输系统的研究与实验,经一年来的试验运行证明,它不失为一种对传统 Internet(计算机 IP 技术+公共电信网)设施的良好补充。

参考文献

- 1 陈彦,徐宏炳,王能斌. WWW 分布数据源研究——数据模型和查询语言. 软件学报, 1998
- 2 李光亚,周学海,等. 一种基于语义网络的开放式超媒体系统结构. 计算机科学, 1997
- 3 李幼平. 用“广存”机制帮助 INTERNET. 计算机世界, 1997

(上接第 66 页)

参考文献

- 1 Ellis C A, et al. Groupware: Some Issues And Experiences. Comm. ACM, 1991, 34(1): 39~58
- 2 Palmer J D, Fields N A. Computer-Supported Cooperative Work. IEEE Computer, 1994, 27(5): 15~17
- 3 Bentley R, et al. Architectural support for cooperative multiuser interfaces. IEEE Computer, 1994, 27(5): 37~45
- 4 Zhao J, Hoppe H U. Supporting flexible communication in heterogeneous multi-user environments. In: Proc. of the Intl. Conf. on Distributed Computing Systems. 1994
- 5 Reinhard W, et al. CSCW Tools: Concepts and Architectures. IEEE Computer, 1994, 27(5): 28~36
- 6 Shirmohammadi S, et al. Applet-Based Telecollaboration: A Network-Centric Approach. IEEE Multimedia, 1998(April-June): 64~73
- 7 Greif I, et al. A Case Study of CES: A Distributed Collaborative Editing System Implemented In Argus. IEEE trans. on software engineering, 1992, 18(9): 827~839
- 8 Mao B, Xie L. An Object-based Model for Prototyping User Interfaces of Cooperative System. ACM Software Engineering Notes, 1997, 22(2): 72~76
- 9 茅兵, 谢立. 基于对象的协同计算模型. 中国科学(E 辑), 1997, 27(6): 542~547