维普资讯 http://www.cqvip.com

在移动计算环境中基于移动代理的缓存失效方案*)

吴 劲 卢显良 任立勇
(电子科技大学计算机学院 成都610054)

A Cache Invalidation Scheme Based on Mobile Agent in Mobile Computing Environments

WU Jing LU Xian-Liang REN Li-Yong
(College of Computer Science & Engineering, UESTC of China, Chengdu 610054)

Abstract Caching can reduce the bandwidth requirement in a mobile computing environment as well as minimize the energy consumption of mobile hosts. To affirm the validity of mobile host' cache content, servers periodically broadcast cache invalidation reports that contain information of data that has been updated. However, as mobile hosts may operate in sleeping mode (disconnected mode), it is possible that some reports may be missed and the clients are forced to discard the entire cache content. In this paper, we present a cache invalidation scheme base on mobile agent in mobile computing environments, which can manage consistency between mobile hosts and servers, to avoid losing cache invalidation reports.

Keywords Cache invalidation, Mobile host, Mobile agent, Mobile support station, Home cache

1 引言

缓存技术是分布式计算环境中的重要技术,它可以改善系统的整体性能(如查询响应时间、吞吐量等),而移动计算的网络环境是一种特殊的分布式环境,与传统的分布式系统相比,它具有鲜明的特点:移动性、断接性、带宽多样性、可伸缩性、弱可靠性、网络通信的非对称性、电源能力局限性等等。这些特点使得缓存技术在移动计算环境中尤为重要,因为缓存能有效减少带宽需求,并能节省移动计算机的能耗。

在移动计算环境中的缓存技术主要考虑两个方面的问题:(1)缓存什么?(2)如何保证数据的一致性?本文讨论的是第二个方面问题,即已经解决了缓存什么问题后,如何保证缓存内容和服务器数据的一致性问题。目前文献中普遍采用的是缓存失效报告法,即服务器定期广播缓存失效报告,该报告中包含了最近被更新的对象,根据报告内容,客户可使缓存中被更新对象失效,以保证缓存的有效性。但是,移动主机可工作在睡眠状态(与基站断连),有些报告有可能丢失,因此不少缓存失效报告方案,都强制客户丢弃整个缓存的内容,或采取一些策略挽救缓存内容。本文提出了一种基于移动代理的缓存失效方案,能有效保证缓存内容和服务器数据的一致性,避免报告丢失带来的不良后果。

2 基于代理的缓存失效方案

2.1 系统模型

本文所考虑的移动计算环境的系统模型如图1所示。在这个环境中,移动主机(Mobile Host,MH),通过无线网络的移动服务支持站(Mobile Support Station,MSS),查询固定网络中的数据库服务器。

移动计算环境中的计算机节点可以分为两类:一类是连接在固定网络上的固定主机,称为基节点;另一类是通过无线网络与基节点相联的移动节点(移动主机)。固定节点上可包含文件服务器和数据库服务器;而移动服务支持站MSS是一类特殊的固定节点,它带有支持无线通信的接口,负责建立一个无线网络单元(无线广播单元或无线局域网单元)。无线网络单元内的移动主机通过无线网络与MSS连接,进而通过MSS和固定网络上的固定节点以及其它移动主机连接。

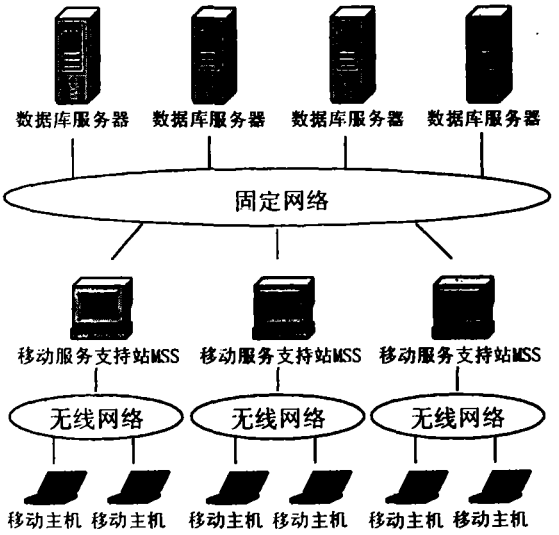

图1 系统模型

一个移动主机有两种状态:清醒状态(Awake Mode)和睡眠状态(Sleeping Mode)。当移动主机处于清醒状态时,它可与服务器相连,接收缓存失效报告;当移动主机与网络断连,不论是主动的还是被迫的,在断连期间,它都接收不到缓存失效报告,所以主机处于睡眠状态。当处于睡眠状态的断连主机重新接入网络,我们称主机被唤醒(Wakeup)。

我们考虑下面的情况:一个运行在移动主机上的客户进程,通过消息与数据库服务器通信,即客户发送一个上行请求(Uplink Request)向数据库服务器请求所需数据,数据库服

*)本文由国防973重大基础研究项目资助,编号:973-1-4-2。吴 劲 讲师,博士生,主要研究方向为计算机网络及分布式数据库技术。卢显良 教授,博士生导师,主要研究方向为计算机网络技术及应用。任立勇 讲师,博士,主要研究方向为计算机网络技术及应用。

维普资讯 http://www.cqvip.com

务器响应请求，通过下行(Downlink)通道发送相应数据。为了减少请求的数量，客户可在本地缓存数据库中的部分数据，这样当缓存中有客户所需数据时，就不需向数据库服务器提出请求，节省了网络带宽。

但当数据库中的数据被更新后，缓存中的部分数据有可能失效，为了保证缓存数据与共享的数据库服务器数据的一致性，当数据库数据被更新后，就向网络广播缓存失效报告。该报告中应包括最近被更新的对象，相应节点接收到报告后，根据报告内容使缓存中被更新的对象失效，以此保证缓存数据的有效性。

最理想的情况就是移动客户可以接收到所有失效报告，那只要服务器定期广播失效报告就可以保证缓存数据的有效性。但实际情况并非如此，移动客户可能经常与网络断连，断连后可能错失报告，即使不断连，由于其它原因也可能造成报文丢失。本文提出的基于移动代理的缓存失效方案，能有效解决这个问题。

2.2 方案介绍

基于移动代理的缓存失效方案的缓存体系结构如图2所示。对于每个移动主机(MH)，都在本机有一个数据缓存区；除此之外，在该主机所在的无线网络内，都有一个移动服务支持站(MSS)，MSS 上有 n 个移动代理(Mobile Agent)负责分别管理该网络内相应的 n 个 MH 的缓存副本，该副本称为本地缓存(Home Cache)。

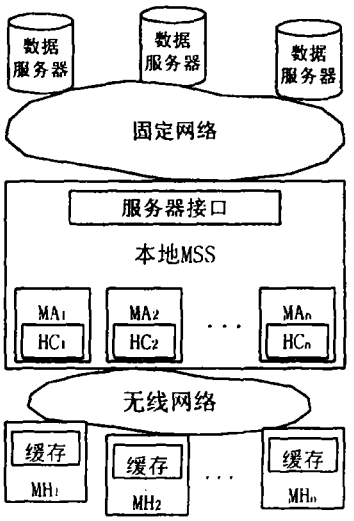

图2 缓存体系结构

本方案基于下面的假设：

1. 当网络上的任一数据库服务器的数据项被更新，都会有相应的失效性报告发送给所有的移动服务支持站(MSS)。

2. 移动主机(MH)能检测其是否联网。

3. 每个 MSS 上都有为移动代理(MA)提供的运行环境——移动代理环境(MAE)

4. 当 MH 上的缓存内容被更新，可通知 MSS 上的相应移动代理(MA)更新相应的本地缓存(HC)。

5. 由 MH 所在无线网络单元的 MSS 上的相应 MA 维护其 HC，并通知 MH 它从服务器接收到的失效通知。

2.2.1 *在特定 MSS 上移动代理的工作原理* 对于一个具体的 MSS 而言，在它的覆盖范围内有 n 个移动主机(MH_i，$1 \leqslant i \leqslant n$)，对于任一个 i，都有一个相应的移动代理 MA_i 维护其 HC_i，此 HC_i 是对应 MH_i 的缓存副本。HC_i 是由一系列的记录清单组成的，每个记录是一个三元组(x，TS，invalid_flag)。其中 x 表示数据项，TS 是服务器在失效报告中提供的时戳(Time-Stamp)，invalid_flag 是失效标记，缺省值为 FALSE。

因为本方案失效报告的传送是异步的，失效报告先缓存在 HC_i 中，当移动主机与 MSS 连接时才由 MA_i 传送报告给 MH。若失效项已发给 MH 但没收到确认，invalid_flag 被标记为 TRUE。

每个移动主机维护本机的经常被访问缓存数据项，在回答任何应用查询之前，它首先检查数据的一致性状态。具体过程如下：当 MSS 收到来自服务器的失效报告，MSS 通过和 MA 协商决定哪些 MH 集需要这些数据并发送给它们。MH 收到失效信息，它使本机的缓存中相应数据项失效。当 MH 从应用层收到查询请求，检查缓存数据的有效性，如果数据项有效，在本地就可以满足查询；否则，向 MSS 中的 MA 发送上行查询，MA 再向服务器提出此请求，当 MA 从服务器收到请求数据后，先增加到 HC 中，再发送给 MH。

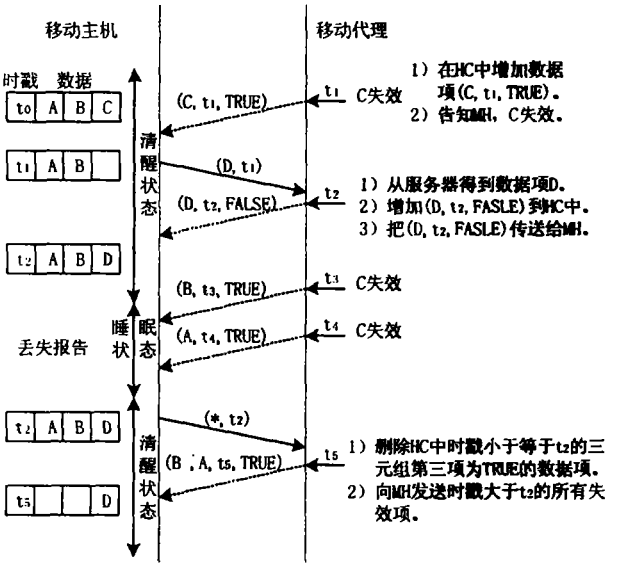

图3 工作原理图

一个移动主机可以工作在睡眠状态和清醒状态，在睡眠状态，移动主机无法收到由 MA 发送的失效报告。我们采用时戳(Time Stamp)的方法，由 MA 异步传送失效报告给 MH，以此避免报告丢失。具体方案如下：当 MA 收到一个来自 MH 带有时戳 t 的消息，它删除在 HC 中所有时戳小于等于 t 的失效报告，然后向 MH 发送所有时戳大于 t 的失效报告。这样，当 MH 从睡眠状态中醒来，它只需向 MA 发送一条带有其缓存时戳的消息 MA 就可以把 t 之后的所有失效报告发送给它。采用该方案，移动主机根据时戳可以判断它断连后哪些数据项被改变了。注意：移动主机断连后，它首先要从 MA 接收所有没有收到的失效报告，在这之前所有查询都会被延迟响应。

下面我们举例说明该方案的工作原理，其原理如图3所示。初始化时移动主机(MH)的时戳为 t_0，MH 的缓存中有三个成员 A、B、C。当移动代理(MA)收到来自服务器的失效报告，通知说在 t_1 时刻 C 被更新，它向 HC 中增加(C，t_1，TRUE)，并通知 MH 数据项 C 失效。MH 收到来自 MA 的失效报告，MH 把其时戳更新为 t_1，并把数据项 C 从缓存中删除。之后，要查询数据项 D，MH 发现缓存中无此项，于是向 MA 提出查询请求。MA 从服务器取得数据项 D 后，先把(D，

维普资讯 http://www.cqvip.com

t_2,FALSE)增加到HC中,再发送给MH。MH收到数据,因最后一项为FALSE,知道这是所求数据,把它写入缓存,更新时戳为t_2,并回答查询。假设此时MH进入睡眠状态(即可能是主动断连,也可能是网络故障),如图3所示,MA发送的失效项B、A被丢失。当MH被唤醒后,它发送一个探测消息(*,t_2)。MA收到此消息后,删除所有时戳小于等于t_2的三元组最后一项为TRUE的数据项,本例删除(C,t_1,TRUE);然后发送所有时戳大于t_2的失效项。

2.2.2 *移动代理的迁移* 前一小节我们讨论了移动代理在特定MSS上的工作原理,但随着移动主机的移动,移动代理也必须跟着迁移,如何迁移是这一小节需要讨论的问题。

移动代理系统由移动代理(MA)和移动代理环境(Mobile Agent Environment)两个部分组成。MAE是分布在网络各计算机设备上的软件系统,它也称为移动代理服务器和移动代理平台,它建立在操作系统之上,为MA提供运行环境。MA是只能存活在MAE中的软件实体,MA的移动便是从一个MAE移动到另一个MAE。本文中的假设3已说明,要求所有的MSS装载MAE。

移动代理的迁移原理如图4所示,其工作原理说明如下:

(1)假设移动主机MH_i从无线网络N_i漫游到无线网络N_j;

(2)MH_i向网络发送一个迁移消息,该消息包含其当前的缓存时戳t;

(3)当在MSS_j上的MA_i收到此消息后,它先在MSS_j上创建一个新的移动代理MA'_i,再把HC_i中所有大于等于t的失效项传给新代理。

(4)新的代理MA'_i被创建后,与Mh_i通信复制缓存副本。

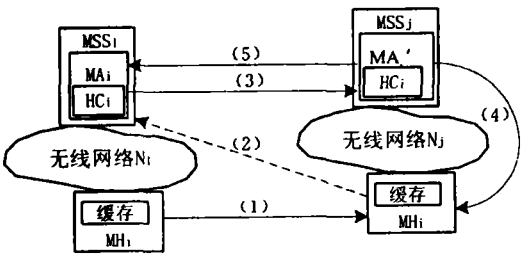

图4 移动代理的迁移

(5)在确认新代理收到所有大于等于t的失效项后,在原位置的MA_i被撤销,在此期间的所有由服务器发送的失效报告由新代理接管。

结束语 本文提出的基于移动代理的缓存失效方案最大的优势在于:不论是什么原因引起的失效报告丢失,例如移动主机主动断连、网络故障引起的断连或单纯的报文丢失,它都能处理,而且方法简单。特别是当MH位于某特定MSS覆盖网络时,工作非常稳定,若MH跨网段移动频繁,其代理迁移方案还有待进一步改进,这将是我们下一步工作的重点。

参考文献

1 Barbara D, Imielinski T. Sleepers and Workaholics: Caching Strategies in Mobile Environments (Extended Version). MOBIDATA: An Interactive J. Mobile Computing, 1994,1(1)

2 Barbara D, Imielinski T. Sleeper and Workaholics: Caching Stragegies in Mobile Enviroments. Very Large Database J. ,Dec. 1995

3 Kahol A, et al. A Strategy to Manage Cache Consistency in a Disconnected Distributed Environment. IEEE transactions on parallel and distributed systems,2001,12(7)

4 Tan K-L, Cai Jun, Ooi B C. A Evaluation of Cache Invalidation Strategy in Wireless Environments. IEEE transactions on parallel and distributed systems, 2001,12(8)

5 Liu G Y,McGuire G Q,Jr. A Mobility-Aware Dynamic Database Caching Scheme for Wirless Mobile Computing and Communication. Distributed and Parallel Database, 1996,4:271～288

6 Sistla A P. Wolfson O,Huang Y. Minimization of Communication Cost Through Caching in Mobile Enviroments. In: Proc. ACM Special Interest Group on Management of Data, May 1994

7 Cai C, Tan K L. Energy-Efficient Selective Cache Invalidaion. Wireless Networks, 1999,5(6):489～502

8 Sistla A P,Wolfson O,Huang Y. Minimization of Communication Cost Through Caching in Mobile Enviroments. IEEE Trans. Parallel and Distributed Systems,1998,9(4):378～389

9 Mobile Agent Facility Specification. OMG TC Document ef/xx-xx-x,1997

10 Wong D, et al. Concordia: An infrastructure for collaborating mobile agents. In: Mobile Agent-First International Workshop, MA'97. Berlin,1997

11 The Agent Society. Agent Product and Research Activities. 2000. http://www.agnent.org

(上接第66页)

结束语 在入侵检测技术中,需要解决的关键问题是未知恶意行为的检测问题。基于数据挖掘的入侵检测方法为检测未知的恶意行为提供了一种可行的方法。这种方法在检测精度、算法效率、实时检测以及系统实现等方面还存在着一些问题,有待于进一步研究和解决。

参考文献

1 Jiawei H, Micheline K 著, 范明等译. 数据挖掘—概念与技术. 机械工业出版社,2001

2 Wenke L, Salvatore J S. Data Mining Approaches for Intrusion Detection. In: Proc. of the 7th USENIX Security Symposium, January, 1998

3 Wenke L, Salvatore J S. Automated Intrusion Detection Using NFR: Methods and Experiences. In: Proc. of the Workshop on Intrusion Detection and Network Monitoring, April, 1999

4 Henry S, Teng K C, Stephen C L. Security Audit Trail Analysis Using Inductively Generated Predictive Rules. In: Proc. of the 11th National Conference on Artificial Intelligence Applications, IEEE, IEEE Service Center, Piscataway, NJ, March 1990. 24～29

5 Debar H, Becker M, Siboni D. A Neural Network Component for an Intrusion Detection System. In:Proc. of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, May 1992

6 Rebecca G B 著, 陈明奇译. 入侵检测. 人民邮电工业出版社, 2001

7 Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI, 1995

8 Matthew G S, Eleazar E, Erez Z, Salvatore J S. Data Mining Methods for Detection of New Malicious Executables. To appear in IEEE Symposium on Security and Privacy, May 2001