

基于 GPRS 网络的强一致性两级缓存系统框架^{*}

谭 劲 朱光喜

(华中科技大学电子与电信工程系 武汉 430074)

摘 要 无线移动环境中缓存的主要目的是减少对无线带宽资源的占用和节省电池能量,然而移动无线终端的漫游与经常断开连接又给缓存内容的一致性带来了一系列新的问题。本文针对目前运行的 GPRS 网络,提出了在用户端(移动终端)和 GPRS 骨干网中添加验证服务器 VS(Validation Server)对数据进行两级缓存的系统框架和缓存强一致性策略。该框架简化了无线移动环境下维护缓存一致性的复杂性,有效地降低了对无线带宽的占用和数据库服务器的负载,支持移动终端断开连接的时间任意和在一个公众陆地移动通信网 PLMN 网内的漫游,具有很强的实用性。

关键词 移动终端,缓存,强一致性,失效报告,移动计算

The System Framework of Caches with Two Levels and Strong Consistency Based on GPRS Networks

TAN Jin ZHU Guang-Xi

(Dept. of Electronic & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074)

Abstract The main purposes of cache are saving radio bandwidth and power energy consumption in mobile wireless environments. However, the roaming and frequent disconnections of the mobile terminal bring a series of new problems to cache consistency. Aiming at the GPRS network running at present, we have proposed the system framework of caches with two levels and strong consistency scheme, one level is at mobile terminal, and another is at VS (Validation Server). The framework and scheme can simplify the complexity of maintaining cache consistency in mobile wireless environments, reduce the radio bandwidth consumption and the workload of database server efficiently, support the arbitrary disconnection intervals of mobile terminal and roaming in a Public Land Mobile Network, and have good practicability.

Keywords Mobile terminal, Cache, Strong consistency, Invalidation report, Mobile computing

1 概述

在无线移动计算环境中,大量的低能耗设备通过无线信道查询数据库,移动客户为了延长电池的使用时间会经常断开连接,并位于不同的信号单元连接到不同的服务器。在移动设备的本地缓冲区中缓存经常访问的数据对象是一种减少查询延迟、节省带宽、降低电池消耗、改善系统性能的有效方法。许多应用程序,如移动银行事务处理和股票交易信息需要移动终端 MT(Mobile Terminal)缓存的数据与服务器端不断更新的数据保持强一致性。然而,MT 的移动(漫游)及有意或无意断开连接又给缓存内容的一致性带来了一系列新的问题,目前对无线移动环境下的缓存研究比较多的集中在维护缓存数据一致性上^[1,5~7]。文[4]最先提出了一种无状态的广播失效报告 IR(Invalidation Report)策略,采用每隔 L 秒周期性地(同步)广播数据库服务器端所有最近更新的数据项标识符 ID (不含数据本身)和更新时的时间戳 TS (Timestamp)组成的数据对队列 (ID_i, TS_i) 的 IR 报告,并采用窗口大小参数 w 来限制 IR 报告的长度。客户端(MT)每次向数据库服务器发出数据请求时,如果请求数据项缓存在缓冲区中,就先监听并收下一个 IR 报告,然后检查缓存数据的有效性。如果数据有效,就直接使用缓存的数据,否则向服务器请求新的数据。尽管基于 IR 的策略保证了缓存数据的强一致性,但主要有以下缺点:

- MT 需要等待并接收下一个 IR 报告才能使用已缓存

的数据,平均等待时间为 $L/2$;

- 如果 MT 断开连接的时间超过了 $w * L$,缓存的数据全无效,降低了缓存的命中率;

- 由于 IR 信息的增加,该机制不适合数据更新很快的大型数据库^[2,7];

- IR 报告中包含了未被缓存而已更新的数据,浪费了广播带宽;

- 所有 MT 收到 IR 报告后聚集向服务器请求更新的数据,形成突发的网络流量。

许多后续研究是基本 IR 策略的改进,文[5]集中在解决 MT 长时间断开问题和利用二进制位缩短 IR 报告的长度,文[6]主要集中在降低 MT 请求数据时的长时间等待和有效利用广播带宽。目前有状态的无线移动环境缓存一致性的研究很少,文[7]提出了一种 AS(Asynchronous Stateful) 维护环境缓存一致性的有状态算法,服务器端维持每个 MT 的 HLC(Home Location Cache)状态,根据 MT 缓存的内容异步地发送 IR 报告。

基于 IR 的策略除了上述主要缺点外,我们认为还有以下两个方面需要加以改进:

(1) 客户端的计算量大

在验证缓存数据的有效性时,MT 需要监听并接收 IR 报告,然后在 IR 报告中查找本地已缓存数据项的 ID ,通过比较 IR 报告中与缓存数据的时间戳来确定缓存的数据是否有效,如果服务器在 $w * L$ 时间内更新的数据项很多(IR 报告很

^{*} 基金项目:国家“863”项目(2001AA123014)。谭 劲 博士后,主要从事网络通信、多媒体技术方面的研究。朱光喜 教授,博导,主要研究方向为数字视音频编码与传输、多媒体通信、宽带无线 IP 通信系统。

长)或客户端缓存的数据没有更新(缓存的数据仍然有效),将会耗费或浪费客户端的处理能力(耗费电池能量)。

(2) 网络中参与维护缓存一致性的节点数多

基于 IR 报告^[2,4~6]中的网络环境是一个抽象的无线网络环境,它假定覆盖每个信号单元(Cell)的 MMS(Mobile Support Station)是完全复制源数据库的数据服务器(对 MT 来说,MMS 就是数据库服务器),MMS 间采用有线连接,并假定在数据库服务器与 MMS 间存在某种数据一致性的机制。当这种方案应用于实际无线网络(如 GPRS)环境时,所有的基站 BS(Base Station)将参与维护数据的一致性,而多节点数据库完全实时复制的部署与开销是很大的。文^[3]也指出了基于 IR 报告的策略能比较好地适用于无线局域网,不适用于像 GPRS 这类的商业网络。

针对目前运行的 GPRS(General Packet Radio Service)网络,我们提出了在客户端(MT)和 GPRS 骨干网中添加验证服务器 VS(Validation Server)对数据进行两级缓存的系统框架和强一致性缓存策略,MT 端缓存数据的主要目的是降低对无线带宽的占用和电池能量消耗;VS 端则是为了减少 Internet 网络流量和降低服务器负载,同时兼有维护 MT 端缓存数据一致性的功能。该方案只使用一个服务器维护缓存数据的一致性,克服了基于 IR 报告策略的所有缺点,支持在一个公众陆地移动通信网 PLMN(Public Land Mobile Network)网内的漫游,具有很强的实用性。

2 系统结构

我们提出的缓存系统结构如图 1 所示。GPRS 是一种基于 GSM 系统的无线分组交换技术,提供端到端的、广域的无线 IP 连接。当 MT 进行 GPRS Attach 时,向 SGSN(Serving GPRS Support Node)发出带有 MT 设备的国际移动用户标识符 IMSI(International Mobile Station Equipment Identity)信息的请求,SGSN 向 HLR(Home Location Register)发出请求确认该 IMSI 设备已授权后,发给 MT 激活接受信息,并在 SGSN 中建立该设备的有效位置信息和路由信息^[9]。只有 GPRS Attach 成功的 MT 才能通过 GGSN(Gateway GPRS Support Node)访问 Internet 上的数据(此时 MT 处于 READY 状态),SGSN 中纪录有 MT 在一个 PLMN 网络中的位置信息,也就是说 MT 是否进入网络是可以确定的。

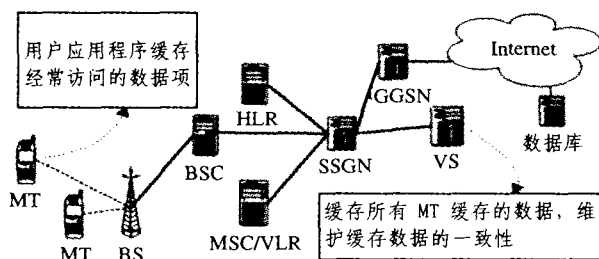


图 1 缓存系统结构

从图 1 中可以看出:缓存系统框架中有三个关键的部分,即位于 Internet 中的数据库、位于 GPRS 骨干网中的 VS 和在 BS 无线信号覆盖范围内的多个 MT。为了便于描述我们提出的缓存系统框架,做如下假定:

(1)Internet 上的数据库是数据项 $d_1, d_2, \dots, d_n$ 的集合,数据项 d_i 是 MT 缓存数据的基本单位,数据的更新只在服务器端进行(MT 只能读取数据项)。

(2)当数据库中的数据项 d_i 更新时,数据库服务器会将该数据项的 ID_i 可靠传送给验证服务器 VS(有线连接)。

(3)在 MT 进入网络(在线)后,其无线连接是可靠的。

(4)当 MT 替换出一个已缓存的对象时,将其 ID 号通知 VS(无线连接)。

(5)VS 有足够的资源在 MT 与数据库之间进行数据的转发和缓存。

假定(1)、(2)、(3)与其它研究^[2,4~7]基本相同,但我们的研究省去了时间戳 TS。

3 数据结构与强一致性算法

缓存系统的主要目标是在保证 MT 端缓存数据与服务器端变化的数据强一致性的条件下,优先降低 MT 对无线带宽的占用和电池能量消耗,其次是减少 Internet 网络流量和降低服务器负载。

3.1 MT 端

MT 向 SGSN 注册后,驻留在 MT 上的应用程序(如浏览器)缓存用户经常访问的数据项。根据 MT 上的存储配置,缓存的数据项可以是一个,也可以是多个,MT 端缓存数据结构如图 2 所示。我们考虑 MT 端采用 LRU(Least Recent Used)缓存替换算法,该算法实现简单,效率很高,很适合 MT 这样资源受限的设备。图中,LRU 排序表示根据 LRU 算法排出的顺序;MT_Mark 表示缓存的数据项是否有效(True 表示有效),其值仅在接收到 VS 传送来的无效通知时改变为 False;Data File 表示已缓存数据项的文件名。当 MT 请求一个数据项时,这里分两种情况:

(1)请求的数据项已缓存:计算该数据项的 LRU 序值,如果 MT_Mark 的值为 True,就直接使用本机已缓存的数据(节省了无线带宽);否则通过 VS 获得该数据项,然后更新已缓存的数据。

(2)请求的数据未缓存:通过 VS 获取请求的数据,将含有数据项的 ID 、MT_Mark 的值(True)、计算的 LRU 序值和缓存数据项文件名的记录插入图 2 的表中,如果缓存前缓冲区已满,先释放 LRU 排序最后的数据项。

Data ID	LRU 排序	MT_Mark	Data File
7	1	True	Data7
1724	2	False	Data1724
...	...	...	...

图 2 MT 端数据结构

3.2 VS 端

VS 端有两个表,如图 3 所示。表 1 以所有 MT 已缓存数据 ID 项为索引,计数器 Counter 表示缓存该项数据的用户数,Data File 表示已缓存数据项的文件名;表 2 以 IMSI 为索引(为了便于说明,拆分成了两个子表),标记 VS_Mark 表示对应的 MT(IMSI)是否需要更新数据(False 表示需要更新),两表通过数据项 ID 值相关联。当 VS 接收到 MT 发出的数据请求时,这里分两种情况:

(1)请求的数据项已缓存:将图 3、表 1 中对应数据项的 Counter 加 1,在表 2 中新增该 MT 的 IMSI,VS Mark 置为 True,同时将缓存的数据项响应给请求的 MT,降低了 Internet 网络流量和服务器负载。

(2)请求的数据项未缓存:先将 MT 的请求转发给数据

库服务器,接收到服务器的响应后,将数据缓存在本地存储设备上;然后将含有数据项的 ID、Counter 的值为 1 及缓存的文件名的记录插入到图 3 表 1 中,同时在表 2 中新增该 MT 的 IMSI(SGSN 传送给 VS),VS Mark 置为 True;最后将缓存的数据项响应给请求的 MT。

表 1			表 2	
Data ID	Counter	Data File	IMSI	VS-Mark
7	3	Data7	IMSI1	TRUE
...	...	...	IMSI2	TRUE
253	1	Data253	IMSI3	FALSE
...	...	...	IMSI	VS-Mark
1724	2	Data1724	IMSI1	TRUE
...	...	...	IMSI2	TRUE

图 3 VS 端数据结构

3.3 一致性算法

VS 除了转发请求和缓存数据外,另一个重要功能就是维护缓存数据的一致性。当数据库服务器上的数据库发生更新时,就与 VS 进行通信,VS 收到更新信息后,检查更新的数据项是否被 MT 缓存。如果没有缓存,不作任何处理,否则先接收数据项将已缓存的数据项更新,然后将图 3、表 2 中所有缓存该项数据的 VS Mark 全部置为 False,最后根据表 1 中缓存该数据项的用户个数,在表 2 循环查询 SGSN 中的注册数据库,将缓存数据已无效的信息发送给在线的 MT,这里分两种情况:

(1)单个 MT 缓存该项数据(Counter=1):VS 将该数据项的 ID 单播发送给 IMSI,并将该 IMSI 对应的 VS Mark 置为 True;MT 收到更新信息后,只需将本地缓冲区中该数据项对应的 MT Mark 置为 False,如果 MT 在后续访问中还要请求该数据项,就向 VS 提出请求,同时用更新的数据替换已无效的数据;否则该数据项有可能被缓存替换策略移出缓冲区。

(2)多个 MT 缓存该项数据(Counter>1):VS 先将在线用户形成一个组,然后将该数据项的 ID 组播发送给在线的 IMSI,组中 IMSI 对应的 VS Mark 置为 True,不在组中的 IMSI 对应的 VS Mark 的值还是为 False(待 MT 进入网络后再发给其无效信息),MT 收到更新信息后的处理和单播(1)相同。

对于那些有意或无意断开连接(不管断开连接的时间长短)的 MT,当再次连接时,会向 SGSN 注册。SGSN 发给 MT 激活接受信息后,将该 MT 的 IMSI 通知 VS。VS 收到通知后,在表 2 中查找该 IMSI 已缓存的数据项,如果对应的 VS Mark 的值为 False,就将其 ID 加入到更新数据项的 ID 列表中,然后将该 ID 列表单播发送给 IMSI,并将该 IMSI 对应的 VS Mark 置为 True,否则不作任何处理。

3.4 移出缓存对象

MT 是一种资源有限设备,它维持一个 LRU 序值,当缓存区满时,就移出 LRU 排序最后的数据项,并将该对象的 ID 通知给 VS,但 VS 端并不一定要移出该数据项。当 VS 收到 MT 端的移出通知后,就从表 2 中清除该 IMSI 缓存该数据项的记录,同时将表 1 中该项数据对应的 Counter 减 1。只有当表 1 中该数据项 Counter 的值等于零时,才从表 1 中清除该项数据,并删除该数据项对应的数据文件,说明整个系统中已

没有用户缓存该项数据了。

3.5 算法

在我们提出的缓存一致性策略中,使用了 6 种消息,表 1 列出了消息名、发送/接收端及其说明。图 4 描述了 VS 端缓存和维护数据一致性算法的伪代码。

表 1 系统中使用的消息

消息名	发送端	接收端	说明
Uncached (ID)	MT	VS	MT 替换出了已缓冲的数据项 ID
Request (ID)	MT	DB Server	MT 向数据库请求数据项 ID
Online (IMSI)	SGSN	VS	IMSI 已在线(进入网络)
Updated (ID)	DB Server	VS	数据库服务器中的数据项 ID 已更新
Invalidated (IDList)	VS	MT	MT 缓存的数据项 ID 列表已无效
IsOnline (IMSI)	VS	SGSN	VS 向 SGSN 查询 IMSI 是否在线

```

Do While(True)
{
  Receive message from MT,DB Server,SGSN;
  Switch(message)
  {
    Case Re quest(ID)
    {
      If ID∈ Table1
      {
        Counter=Counter+1;
        Insert a new item in Table2 and Set VS-Mark= True;
      }
      Else
      {
        Request data item ID from DB server and cache it;
        Insert a new item in Table 1 and Set Counter=1;
        Insert a new item in Table 2 and VS-Mark= True;
      }
    }
    Case Uncached (ID)
    {
      Counter=Counter-1;
      If Counter=0 Delete data ID from Table 1;
    }
    Case Invalidated (ID)
    {
      For each IMSI that data ID is cached in Table 2
      {
        Set VS-Mark=False;
      }
      If Counter=1
      {
        Unicast Invalidated (ID) to IMSI and Set corresponding VS-Mark= True;
        Break;
      }
      //形成在线用户的列表
      Send IsOnline (IMSI) to SGSN and Generate a List of IMSI which is online;
      //利用用户访问的局域性向 MT 组播无效 ID
      GroupCast Invalidated (IDList) to the List and Set corresponding VS-Mark= True;
    }
    Case Online (IMSI)
    {
      For each data ID that IMSI Cached Generate a List of ID which is updated;
      //如果 VS-Mark=False,向 IMSI 单播无效 ID
      If VS-Mark=False
      {
        UniCast Invalidated (IDList) to IMSI and Set corresponding VS-Mark= True;
      }
    }
  }
}

```

图 4 VS 缓存和维护数据一致性算法

4 性能分析

我们提出基于 GPRS 网络的强一致性两级缓存系统框架的主要作用是降低对无线带宽的占用和电池能量消耗,减少 Internet 网络流量和降低服务器负载,同时还维护缓存数据的一致性。

4.1 节省带宽资源

当 MT 端发出数据请求时,如果在本地缓冲区缓存命中,就不用将请求转发给 VS,节省了无线带宽;如果未在 MT 端缓存命中的请求在 VS 端发生缓存命中,就不用将请求转发给数据库服务器,节省了有线带宽。只有在 MT 和 VS 端都没有缓存命中的请求,才通过无线(MT 到 VS)和有线(VS 到服务器)两种信道传送到源服务器。假定在 t 时间内,客户端处于 READY 状态的 MT 总数为 M ,所有 MT 端缓存的平均命中率为 H_1 ,MT 对数据项的请求是随机的,其请求到达速率为 λ ,VS 端缓存的命中率为 H_2 ,数据项的大小为 B 字节,用 BW_1 、 BW_2 分别表示本文提出的框架系统节省的无线和有线带宽,则:

$$BW_1 = M * (1 - e^{-\lambda}) * H_1 * B \quad (1)$$

$$BW_2 = BW_1 + M * (1 - e^{-\lambda}) * (1 - H_1) * H_2 * B \quad (2)$$

从式(1)、(2)可以看出,MT 缓存的数据项个数越多,命中率 H_1 越大,节省的无线和有线带宽也越多; H_2 的大小决定于 MT 访问数据项的局域性,其值越大,节省的有线带宽就越多。

4.2 维护数据一致性

目前维护无线移动环境下缓存数据一致性主要有同步的 IR^[4~6]和异步的 AS^[7],在不考虑数据源服务器与 MSS 或 VS 交换数据更新信息的开销的情况下,VS 与 IR、AS 策略定性比较如表 2 所示。此外,MT 的漫游也是移动环境下维护缓存一致性需要考虑的一个重要因素,IR、AS 都是假定所有节点间数据的完全复制来支持 MT 的移动性,即 IR 策略中所有 MSS 广播相同的 IR,AS 策略中在 MSS 间复制相同的 HLC,而 VS 自动支持 MT 在一个公众陆地移动通信网 PLMN 中的漫游(见图 1),不需要另外的开销。

表 2 IR、AS、VS 参数对比

比较的参数	IR	AS	VS
MT 断开连接的最长时间	$W * L$	任意	任意
上行请求次数*	0	1	1
访问已缓存数据的延迟	0 to L	0	0
服务器端是否传送未缓存更新数据的 ID	Yes	No	No
是否向未进入网络的 MT 传送更新数据 ID	—	Yes	No
是否形成突发的网络流量	Yes	No	No
MT 获取已更新数据 ID 的计算量	大	小	小
参与维护数据一致性的节点数	多个	多个	1 个
MT 端缓存一个数据项需要的位数(不包括数据本身)**	64	67	33

*——AS 在断开后再次连接后有一次上行传输;VS 在已缓存的对象

被移出时有一次。

**——假定 ID、TS 各占 4 个字节,标记位占 1 个 Bit 位

结束语 从微处理器芯片到操作系统,从 Web 访问到分布式文件系统,缓存技术一直用于降低访问延迟、节省带宽和提高系统的整体性能。本文针对当前运行的 GPRS 网络和即将来临的 3G 网络,提出了客户端和验证服务器 VS 端两层缓存系统框架,有效地降低了大无线移动用户查询数据库环境对无线和有线带宽的占用,并用异方式解决无线环境下缓存一致性问题,有效地减少了参与维护缓存一致性的节点数和 MT 端的计算量,克服了基于 IR 报告策略的所有缺点,并通过(1)VS 只传送 MT 已缓存且已更新的数据项的;(2)VS 已向已连接进入网络(在线)的 MT 传送信息;(3)取消时间戳降低传送的信息量;(4)利用用户访问数据的局域性面向组内用户进行广播四种方式来减少异步传输的次数,具有较强的实用性。

参考文献

- Barbara D. Mobile Computing and Database——A Survey. IEEE Transactions on Knowledge And Data Engineering, 1999, 11(1):108~117
- Wang Z, Das S, Kumar M. SACCs: Scalable Asynchronous Cache Consistency Scheme for Mobile Environments. In: Proc. 23rd Intl. Conf. on Distributed Computing Systems Workshops, Providence, RI, USA, May 2003
- Lin Y B, Lai W R, Chen J J. Effects of Cache Mechanism on Wireless Data Access. IEEE Transactions on Wireless Communications, 2003, 2(6):1247~1257
- Barbara D, Imielinski T. Sleepers and Workaholics: Caching in Mobile Distributed Environments. In: Proc. of 1994 ACM-SIGMOD Intl. Conf. on Management of Data, Minneapolis, MN, USA, June 1994. 1~2
- Jing J, Elmagarmid A, Helal A, et al. Bit-Sequences: An Adaptive Cache Invalidation method in Mobile Client/Server Environment. ACM-Baltzer Journal on Special Topics in Mobile Networks and Applications, 1997, 2(2):115~127
- Cao G H. A Scalable Low-Latency Cache Invalidation Strategy for Mobile Environments. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(5):1251~1265
- Kahol A, Khurana S. A Strategy to Manage Cache Consistency in a Disconnected Distributed Environment. IEEE Trans. on Parallel and Distributed Systems, 2001, 12(7):686~700
- Li Y J, Chiang C Y, Liu M T. Effective Web Caching for GPRS Networks. In: Proc. of the Intl. Conf. on Computer Networks and Mobile Computing, Beijing, China, Oct. 2001
- Bettstetter C, Vogel H, Eberspacher J. GSM Phase 2+ General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface. IEEE Communications Survey, 1999, 2(3):2~14

(上接第 33 页)

- Kohler E, Morris R, Chen B J, et al. The Click Modular Router. ACM Trans. on Computer Systems, 2000, 18(3):263~297
- Decasper D, Dittia Z, Parulkar G, et al. Router plugins: A software architecture for next-generation routers. IEEE/ACM Trans. on Networking, 2000, 8(1)
- Karlin S, Peterson L. VERA: An Extensible Router Architecture. In: Proc. of the 4th Intl. Conf. on Open Architectures and Network Programming (OPENARCH), April 2001. 3~14
- Peterson L, Gottlieb Y, Hibler M, et al. An OS Interface for Active Routers. IEEE Journal on Selected Areas in Communications, 2001, 19(3):473~487
- Keshav S, Sharma R. Issues and trends in router design. IEEE Communications Magazine, 1998, 36(5):144~151
- Partridge C, et al. A 50Gb/s IP Router. IEEE/ACM Trans. on Networking, 1998, 6(3):237~248

- Semeria C. Internet Backbone Router and Evolving Internet Design. White Paper, Juniper Networks, September 1999. <http://www.juniper.net>
- 徐恪, 吴建平, 江勇, 徐明伟. 通用路由器软件体系结构研究综述. 小型微型计算机系统, 2001(4)
- IBM Corp. PowerNP NP4GS3 Datasheet. <http://www.ibm.com/chips>
- Intel Corp. Intel IXP1200 Network Processor Family. <http://developer.intel.com>
- EZchip Technologies Inc. NP-1 Network Processor Product Brief. <http://www.ezchip.com>
- 徐恪, 吴建平, 吴剑, 徐明伟. 高性能路由器操作系统 HEROS 中的高可用性设计. 863 学术年会, 2001
- Montz A B, et al. Scout: A Communication-Oriented Operation System. In: Proc. of the Fifth HotOS, May 1995. 58~61