



(12) 发明专利

(10) 授权公告号 CN 101539911 B

(45) 授权公告日 2013. 05. 29

(21) 申请号 200810043178. 9

US 7343412 B1, 2008. 03. 11, 全文.

(22) 申请日 2008. 03. 18

祝小华 等. 动态网页加速技术. 《计算机应用》. 2002, 第 22 卷 (第 4 期), 18-21.

(73) 专利权人 盛大计算机 (上海) 有限公司

审查员 李青

地址 201203 上海市浦东新区郭守敬路 356
号

(72) 发明人 张静盛 张在伟

(74) 专利代理机构 上海浦一知识产权代理有限

公司 31211

代理人 丁纪铁

(51) Int. Cl.

G06F 17/30 (2006. 01)

H04L 29/06 (2006. 01)

(56) 对比文件

US 7194506 B1, 2007. 03. 20, 全文.

CN 1937503 A, 2007. 03. 28, 全文.

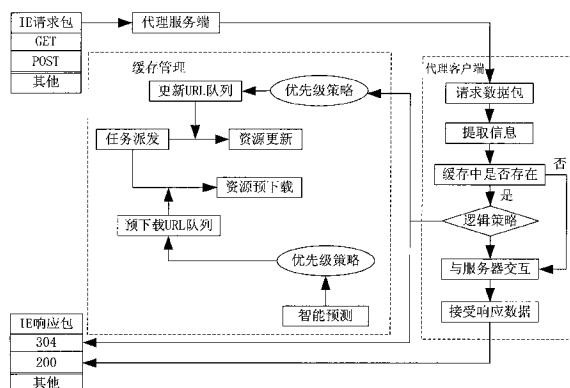
权利要求书1页 说明书4页 附图2页

(54) 发明名称

加快浏览器网页显示的装置及方法

(57) 摘要

本发明公开了一种加快浏览器网页显示的装置包括：代理服务端、代理客户端、缓存管理器。本发明加快浏览器显示的方法，包括：浏览器经代理服务端将 IE 请求包发送至代理客户端，代理客户端分析链接信息对应的网页是否存在于缓存中，不在则经远程服务器显示页面，存在则判断网页是否失效，失效则经远程服务器显示页面，未失效则发送该网页是最新的 IE 响应包，同时将该页面的链接加到缓存管理的更新队列中，进一步分析网页中各页面元素，对于不是静态文件的页面元素发送更新的 IE 请求包。本发明采用显示与更新分离的方法，同时模拟远程服务器发送 IE 响应包，减少了浏览器与远程服务器之间数据的交互，加快了浏览器网页显示的速度。



1. 一种加快浏览器网页显示的装置 ;其特征在于,包括 :

代理服务端,用于接收浏览器的 IE 请求包,并将该 IE 请求包发送给代理客户端 ;

代理客户端,用于接收代理服务端发送的 IE 请求包,提取 IE 请求包中的网络链接信息,分析这个链接所对应的网页是否存在于缓存管理器,若存在,则直接模拟远程服务器向浏览器发送 IE 响应包,然后把这个链接添加到缓存管理器以供浏览器提取并显示该链接对应的网页数据,同时分析网页中的各个页面元素是否为静态文件,对于不是静态文件的页面元素向远程服务器发送该页面元素需要更新的 IE 请求包 ;若不存在,则向远程服务器发送网络链接信息,并向浏览器发送远程服务器返回的 IE 响应包 ;

缓存管理器,用于根据代理客户端的分析,将缓存中的网页进行更新,供浏览器显示。

2. 如权利要求 1 所述的加快浏览器网页显示的装置,其特征在于,该装置还包括智能预测模块,用于根据用户的行为预测用户将要链接的网页,并将该链接信息发送至缓存管理器。

3. 如权利要求 2 所述的加快浏览器网页显示的装置,其特征在于,所述的缓存管理器包括二个 URL 队列,一个是基于代理客户端对页面元素的分析结果的更新 URL 队列,另一个是基于智能预测模块分析的预下载 URL 队列。

4. 一种加快浏览器网页显示的方法 ;其特征在于,包括如下步骤 :

(1) 浏览器发送 IE 请求包至代理服务端 ;

(2) 代理服务端将所述的 IE 请求包发送至代理客户端 ;

(3) 代理客户端分析所述 IE 请求包中的链接信息对应的网页是否存在于缓存中,如果否则向远程服务器发送网页的 IE 请求包,并接收响应数据将该网页的 IE 响应包返回浏览器,结束程序 ;

(4) 对于缓存中存在的网页,代理客户端直接模拟远程服务器向浏览器发送 304 号响应包,并把这个链接添加到缓存管理器,当浏览器接收到 304 号响应时,直接从缓存中取得页面数据并显示 ;

(5) 代理客户端分析所述网页中的各个页面元素是否为静态文件,对于不是静态文件的页面元素向远程服务器发送该页面元素需要更新的 IE 请求包 ;

(6) 缓存管理器根据代理客户端的分析,将缓存中的网页进行更新,供浏览器显示。

5. 如权利要求 4 所述的加快浏览器网页显示的方法,其特征在于,步骤 (5) 所述的静态文件包括静态图片文件和静态脚本文件。

6. 如权利要求 4 所述的加快浏览器网页显示的方法,其特征在于,步骤 (5) 后还包括 :进一步分析所述网页中各个可链接元素,根据用户的行为预测用户可能链接的网页链接,将该链接加入缓存管理器的预下载 URL 队列。

加快浏览器网页显示的装置及方法

技术领域

[0001] 本发明涉及一种网络信息浏览的装置,尤其涉及一种加快浏览器网页显示的装置。本发明还涉及一种加快浏览器网页显示的方法。

背景技术

[0002] 使用浏览器访问互联网,是计算机网络技术的一种最基本应用。目前浏览器的种类繁多 IE、firefox、NetScape、Opera、Maxthon 等,但都受两个因素的影响:访问速度和实时性。访问速度要求浏览器能够在最短的时间内响应用户的 URL 请求,实时性则要求每次把最新的内容呈现给用户。

[0003] 浏览器有一个很重要的部分就是缓存,如何使用缓存机制是评价一款浏览器性能的重要指标。缓存机制就是把浏览器所访问过的页面保存下来,以便在下一次访问同一页面时能够更快地显示。有两种极端的情况,一是只要缓存中有的就不再向服务请求,而是直接提取缓存中的页面并显示,这种情况速度最快但实时性最差。二是不管缓存中是否存在目标页面,都向远程服务器请求目标页面,这样的速度是最慢的,但实时性最高。因而不同程度地使用缓存机制就能在浏览器的访问速度和实时性之间进行调节。

[0004] 这就是浏览器的两个基本问题:数据的显示和更新。例如 IE 是显示更新结合在一起,只有在确保数据最新的情况下才显示,因而就算缓存中的页面已经存在,速度还是会比较慢。例如 Opera 则把数据的显示和更新分离,如果用户请求的页面在缓存中存在,那么就直接显示缓存文件,然后每隔一定时间例如 5 分钟更新一次缓存数据。

[0005] 要想浏览器能有更快速度,就需要采用数据的显示与更新分离的方式。对于这个技术的最好体现就是搜索引擎如 google、baidu 等,当用户请求搜索的时候直接从自己数据库中把已处理好的数据返回给用户,而数据的更新则是在后台通过网络蜘蛛不断地获取页面、处理数据并把保存到数据库中。

[0006] 在一般情况下,浏览器在访问资源时首先进行本地验证,检测资源是否存在且是否过期,如果存在其有效,那么就直接取缓存数据并显示到浏览器界面,这种情况显示的速度是最快的。如果资源存在但已过期,那么就会构造这个资源的验证包到远程服务器去验证。如果远程服务器上的该资源没有更新那么就直接返回 304 包,表示该资源已经是最新,不需要再重新下载;如果远程服务器上的该资源已经更新,那么就返回包含最新资源的 200 包,表示该资源已过期需要重新下载。这个流程如图 1 所示,图 1 中虚线框以内为普通浏览器内核流程简化图,输入为资源请求,一般为 URL 形式,输出为显示到浏览器。

[0007] 中国发明专利公开说明书 CN101075236A(公开日为 2007 年 11 月 21 日)公开了一种加快浏览器网页显示的装置和方法,通过预测当前页面中的链接,预下载各链接的网页保存到本地缓存,从而加快浏览器的浏览速度,上述技术虽然将显示和更新分开,但是对于浏览器每次显示本地缓存的页面时,按常规的操作仍需要与远程服务器建立通讯,收到远程服务器返回的确认当前页面为最新页面的 IE 响应包 304,浏览器才将本地缓存的页面显示,这个交互的过程有时候很快有时候很慢,速度受网络综合因素的影响。且上述技术的

更新根据每次页面的所有链接进行预下载，则下载更新的队列变得很多，很多也并不是用户会去访问的，下载很多多余的网页浪费了带宽也浪费了缓存的空间。

发明内容

[0008] 本发明所要解决的技术问题是提供一种加快浏览器网页显示的装置，它可以加快浏览器的显示速度。为此，本发明还要提供一种加快浏览器网页显示的方法。

[0009] 为解决上述技术问题，本发明的加快浏览器网页显示的装置包括：

[0010] 代理服务端，用于接收浏览器的 IE 请求包，向代理客户端发送网络链接信息；

[0011] 代理客户端，用于接收代理服务端传递的网络链接信息，分析该网络链接信息所对应的网页是否存在于缓存管理器，若存在，则直接模拟远程服务器向浏览器发送 IE 响应包，并分析网页中的各个页面元素是否为静态文件，对于不是静态文件的页面元素向远程服务器发送该页面元素需要更新的 IE 请求包；若不存在，则向远程服务器发送网络链接信息，并向浏览器发送远程服务器返回的 IE 响应包；

[0012] 缓存管理器，用于更新缓存中的网页。

[0013] 本发明利用上述装置加快浏览器网页显示的方法，包括如下步骤：

[0014] (1) 浏览器发送 IE 请求包至代理服务端；

[0015] (2) 代理服务端将 IE 请求包发送至代理客户端；

[0016] (3) 代理客户端分析 IE 请求包中的链接信息对应的网页是否存在于缓存中，如果否则向远程服务器发送网页的 IE 请求包，并接收响应数据将该网页的 IE 响应包返回浏览器，结束程序；

[0017] (4) 对于缓存中存在的网页，直接模拟远程服务器向浏览器发送 304 号响应包，并把这个链接添加到缓存管理器，当浏览器接收到 304 号响应时，直接从缓存中取得页面数据并显示；

[0018] (5) 分析网页中的各个页面元素是否为静态文件，对于不是静态文件的页面元素向远程服务器发送该页面元素需要更新的 IE 请求包。

[0019] 因为本发明用显示与更新分离的方法，同时模拟远程服务器发送 IE 响应包，减少了浏览器与远程服务器之间数据的交互，从而加快了浏览器网页显示的速度。另外本发明还通过预测用户的行为，更新缓存的网页，从而也节省了网页开启的时间，提高了效率。

附图说明

[0020] 下面结合附图和具体实施方式对本发明作进一步详细说明。

[0021] 图 1 是普通情况下浏览器数据处理流程；

[0022] 图 2 是本发明的加快浏览器网页显示的装置的总体结构图；

[0023] 图 3 是本发明的加快浏览器网页显示的方法的流程图；

[0024] 图 4 是本发明用于多个浏览器的工作示意图。

具体实施方式

[0025] 如图 2 所示，本发明的加快浏览器网页显示的装置包括：代理服务端、代理客户端和缓存管理器。其中代理服务端用于接收浏览器的 IE 请求包，向代理客户端发送网络链接

信息。代理客户端,一方面用于接收代理服务端传递的网络链接信息,分析该网络链接信息所对应的网页是否存在于缓存管理器,存在则对网页进行分析,不存在则向远程服务器发送网络链接信息,另一方面向浏览器发送远程服务器返回的 IE 响应包,或模拟远程服务器向浏览器发送 IE 响应包。缓存管理器,用于根据代理客户端的分析,将缓存中的网页进行更新,供浏览器显示。

[0026] 本发明的加快浏览器网页显示的装置还包括智能预测模块,用于根据用户的行为预测用户将要链接的网页,并将该链接信息发送至缓存管理器。

[0027] 如图 3 所示,本发明利用上述装置加快浏览器网页显示的方法,包括如下步骤:

[0028] 首先浏览器发送 IE 请求包至代理服务端,代理服务端将 IE 请求包发送至代理客户端。

[0029] 然后代理客户端提取信息,分析 IE 请求包中的链接信息对应的网页是否存在于缓存中,如果否,则向远程服务器发送网页的 IE 请求包,并接收响应数据将该网页的 IE 响应包返回浏览器,结束程序。

[0030] 对缓存中存储的该网页进行逻辑策略分析,如果网页存在于缓存中,不需要检测是否过期,也不需要经过服务器验证,而是直接模拟远程服务器向浏览器发送 304 号响应包,然后把这个链接添加到缓存管理器,当浏览器接收到 304 号响应时,就直接从缓存中取得页面数据并显示。

[0031] 进一步分析网页中的各个页面元素是否为静态文件(包括静态图片和静态脚本),对于不是静态文件的页面元素向远程服务器发送该页面元素需要更新的 IE 请求包。

[0032] 缓存管理器负责资源的更新和预下载,其中需维护两个 URL 队列,一是更新 URL 队列,二是预下载 URL 队列。在提供代理服务过程,对请求进行加速时,会产生一个资源更新请求,通过优先级测试策略添加到更新 URL 队列。可以分析用户行为,预测将要访问的资源,通过优先级策略添加到预下载 URL 队列。这里优先级策略一般可以直接使用 FIFO 先来后到的方法。之所以需要缓存管理器,就是为了实现显示与更新分离,达到异步的效果。

[0033] 智能预测模块的主要功能是分析用户行为,从而预测用户将要访问的资源。这是一个可扩展模块,可以根据具体需求逐步添加。可以通过打开页面后分析网页中各个可链接元素,根据用户的行为预测用户可能链接的网页链接,将该链接加入缓存管理的预下载 URL 队列。这是由一系列预测方法构成的策略集合,例如当前访问的 URL 是 <http://www.abc.com/xyz>,推测用户很有可能会访问主页 <http://www.abc.com>;例如当用户鼠标在浏览器某个位置上停留一定时间假设为 500ms,这就表示用户很有可能会点击该位置的链接。这些都是可预测的行为。哪怕只提前几百、几十个毫秒都能实现一定的加速效果。

[0034] 图 4 是本发明用于多个浏览器的工作示意图,本发明可以通过 3 个步骤来实现:(1) 开发注入(HOOK) 模块,实现定向连接和智能预测。(2) 如果需要加速的目标程序是由自己开发,那么直接使用 Load Library 加载 Hook 模块即可。否则就开发一个加载器启动浏览器并实现模块注入。(3) 实现加速服务,需要开发代理服务器。浏览器中使用的协议有 HTTP、Secure、FTP、Gopher、Socks 等,但并不需要完全实现所有的协议代理服务。一般情况下,浏览器的最主要功能是实现网页浏览器,此时可以只实现一个 HTTP 代理服务。

[0035] 本发明的加快浏览器网页显示的方法主要有以下几个优势:(1) 加速服务本身作为一个独立的公共加速平台,而不依附于某个具体的浏览器。只要向浏览器注入了实现定

向连接和智能预测功能的模块，该浏览器中的所有请求和预测信息都会被转发到加速服务模块。(2) 不影响浏览器进程的效率。定向模块的功能非常小，几乎所有的逻辑策略都在加速服务模块。(3) 可扩展性强。可以方便地向加速服务模块增加扩展地加速功能。(4) 独立性强。用户可以同时使用具体加速特性的浏览器和普通浏览器，不必修改浏览器配置。

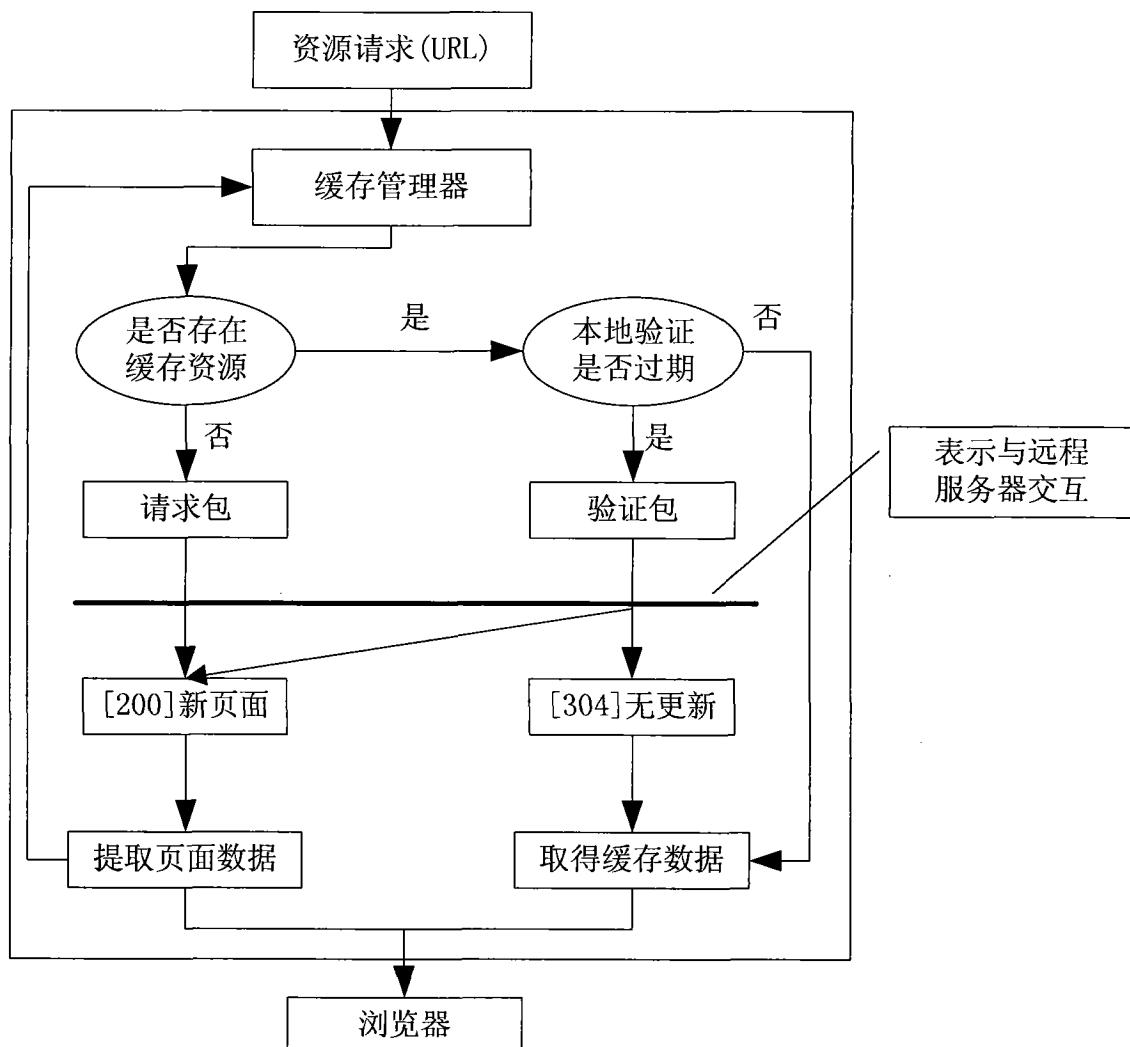


图 1

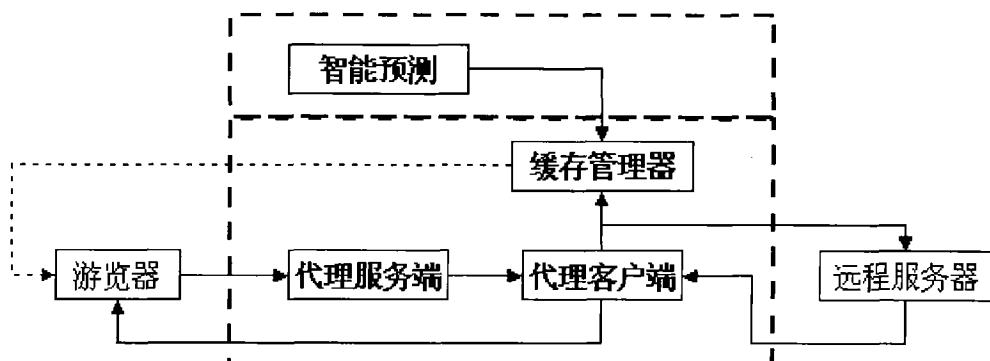


图 2

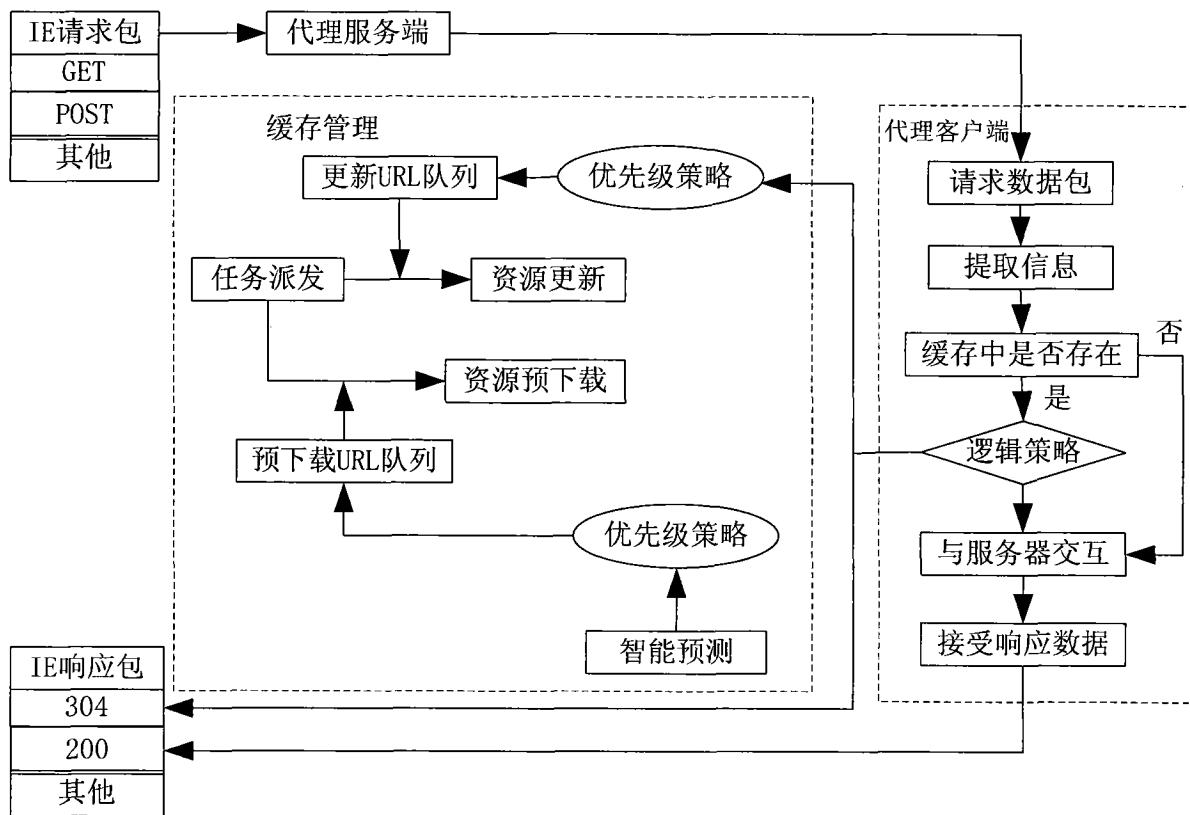


图 3

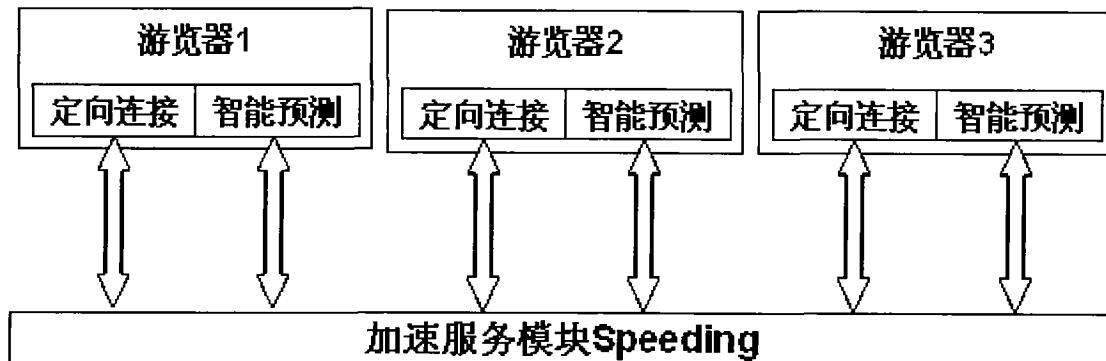


图 4