

# 基于 WebSocket 的缓存式 Web 应用框架设计

## Design for Cache Web Application Framework Based on the WebSocket

宁 君<sup>1</sup>, 张 蓉<sup>2</sup>, 梁正安<sup>1</sup>

NING Jun<sup>1</sup>, ZHANG Rong<sup>2</sup>, LIANG Zheng-an<sup>1</sup>

(1. 广西计算中心, 广西南宁 530022; 2. 北京大学南宁附属实验学校, 广西南宁 530022)

(1. Computing Center of Guangxi, Nanning, Guangxi, 530022, China; 2. Nanning Experimental School Attached to Peking University, Nanning, Guangxi, 530022, China)

**摘要:**【目的】研究 WebSocket 的特性,设计缓存式的 Web 应用框架。【方法】分析划出框架各层次的职责,采用 Javascript 及 memcached 实现实时的缓存读写机制,对框架开发的系统进行指标测试。【结果】在系统测试中检测缓存的命中率,当用户的操作浏览数据达到一定的比例之后,缓存命中率能够维持在一个较高的水平上。在这种情况下,服务器的响应速度、网络吞吐量等都得到极大的改善。【结论】基于框架开发的系统实时性高、负载低、伸缩性好,降低系统的开发成本与难度。

**关键词:** WebSocket Web 缓存 应用框架

**中图分类号:** TP311.5 **文献标识码:** A **文章编号:** 1002-7378(2014)01-0032-03

**Abstract:**【Objective】The characteristics of WebSocket are studied and a cached Web application framework is designed.【Method】The function of every layer in the framework is analyzed and defined. Then indicators of the system developed under the framework are tested by applying a real-time cache read/write mechanism of Javascript and memcached.【Result】When the browsed data reached a certain proportion, the cache hit rate could stay in a higher level during the system test, which greatly improved the server's respond speed and network throughput.【Conclusion】System developed under the framework shows the characteristics of high real-time control, low load, high scalability, low developing cost and difficulties.

**Key words:** WebSocket, Web cache, application framework

【研究意义】在 web2.0 时代,多数应用都是将数据保存在数据库中,应用服务器从数据库中读取数据并显示在浏览器中。但是随着数据量的增加,访问量的提升,往往会出现数据库性能下降,响应缓慢等问题。Web 缓存技术是减轻服务器负载、降低网络拥塞、增强 Web 应用系统可扩展性的有效途径之一,在采用 Web 缓存技术的同时利用 WebSocket 实时、高效的特点在客户端、应用服务器、缓存服务

器、数据库服务器中实现高效实时的数据“推拉”,对提高系统访问效率,减轻服务端负载压力、减少网络带宽消耗,最大限度的挖掘网络及计算资源的潜力有重大意义。【前人研究进展】目前国内外均有基于 WebSocket 的各种技术方案,但应用于框架与缓存的文章尚未见有相关报道。【本研究切入点】深入了解 WebSocket 的特性,研究提出能适用于 Web 框架设计的技术方案,实现框架中 Web 缓存的双向存取。【拟解决的关键问题】研究 WebSocket 特性,利用 WebSocket 实现资源在客户端、服务端、数据端的“推拉”。

收稿日期:2013-11-24

修回日期:2013-12-15

作者简介:宁 君(1984-),男,助理工程师,主要从事软件工程、云计算、平台应用研究。

## 1 框架结构设计技术路线

HTTP 协议是 Web 实时通信的基础,浏览器与服务端建立连接浪费了大量时间和网络吞吐量,加重了服务端负担<sup>[1]</sup>。一般情况下,通过浏览器访问一个网页,需要浏览器发送一个 HTTP Request,服务器接收到浏览器的请求,返回相应的消息。这是大多数的 Web 应用系统数据访问的数据“拉取”模式。服务器端处理客户端的请求需要花大部分的时间来解析 HTTP 标头信息,然而 HTTP request 的标头信息非常长,里面包含的数据则可能只是一个很小的值,却占用很多的带宽。

HTML5 规范中定义,WebSocket 在浏览器和服务端之间采用单 socket 全双工(或者叫双向)传输来“推送”和“拉取”信息。WebSocket 是用于实现服务器与浏览器双向通信的一种解决方案,用于取代一些传统的数据推送方案。服务器与客户端之间交换的标头信息很小,大概只有 2 字节。WebSocket 可以实现客户端和服务端只需进行一次握手建立连接,浏览器和服务端之间就形成了一条快速通道,两者之间就可以直接进行多次数据请求响应。从而摒弃了传统的 HTTP Request 每次的数据请求都需要进行握手以建立连接的过程,网络环境下减少了标头信息的解析过程,大大提高了数据交换的效率(见图 1)。

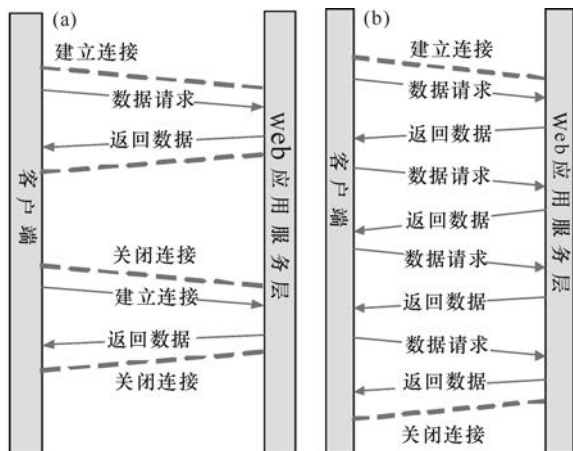


图 1 数据响应的传统处理与 WebSocket 处理流程比较  
(a)传统 Http Reuest 处理方式;(b)WebSocket 处理方式

单数据库的 Web 应用系统在面对高并发高访问的情况下,因为数据连接峰值的到来而接近崩溃边缘,并且在大部分的并发访问中对相同数据的请求浪费了大部分的数据库连接资源。如图 2 所示,我们在 Web 应用系统引入 Web 缓存,以达到减少网络带宽消耗、降低服务器压力、减少网络延迟,加

快页面打开速度等目的。Web 缓存层的存取媒介是内存,数据的读写效率要优于磁盘,采用分布式的缓存层设计,在扩展缓存容量的同时增强灵活性。

为了最小化数据库的负载压力,数据库层采用主从结构模式,用 slave 数据库来完成读取操作,而 master 数据库只负责更新数据、同步复制 slave 数据库、更新缓存。要特别说明的是,在 master 数据更新时,WebSocket 把数据库的更新信息主动的推送至缓存层,这样可以大大减少 Web 层的逻辑复杂度与请求压力。并且采用此种模式能缓解数据库压力,确保数据在异常情况下能够安全备份。

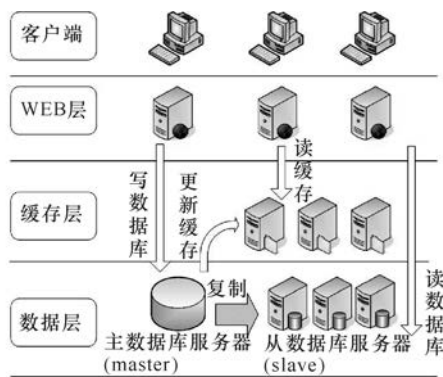


图 2 系统结构模型图

## 2 应用框架设计

框架是一个可复用的设计构件,它规定了应用的体系结构,阐明了整个设计、协作构件之间的依赖关系、责任分配和控制流程,表现为一组抽象类及其实例之间协作的方法,它为构件复用提供了上下文(Context)关系。软件的设计重用性和系统的可扩充性是应用框架设计的重点<sup>[2]</sup>,好的应用框架设计能缩短应用软件系统的开发周期,提高开发质量。与传统的基于类库的面向对象重用技术比较,应用框架更侧重于面向专业领域的软件重用。基于 WebSocket 的缓存式应用框架具有领域相关性,划分为客户层、服务层与数据层(见图 3),开发人员可根据框架进行复合而生成可运行的系统。

客户层负责人与机器的交互之外,通过 Javascript 代码形式的 WebSocket 接口对服务端的数据进行请求。Javascript 能够被所有浏览器支持,客户层的设计能够脱离各种浏览器的限制,轻松实现 WebSocket API 封装的客户层。

服务层的设计由 WebSocket Server 与缓存模块组成。WebSocket Server 构建监听器、消息分发、通道管理、数据处理组件、心跳包与日志管理。监听器用于监听客户请求并过滤请求,根据请求类型交由

消息分发处理。消息分发负责查找实体类,调用缓存或者数据库返回信息,返回消息后通过通道返回至客户端,同时数据库的更新信息也通过 WebSocket Server 的监听器实现对缓存数据的更新。通道是高度有效的通信模块,允许多个模块在应用程序上交互,通道逻辑上是独立的。数据处理组件根据请求处理 JSON、XML 等数据。心跳包是周期性的在客户端与服务端之间发送心跳信息,防止客户端与服务端的握手后,在不传送数据时告知双方的连接状态。日志管理将客户端与服务端的请求与处理信息进行管理。缓存模块实现一个统一的访问组件,访问组件在缓存池中调用缓存存取服务,缓存池中根据调用的 Key 得到经过 hash 算法的 hasy key 访问不同服务器的缓存数据。缓存池还管理着缓存的策略与数据存取方式。

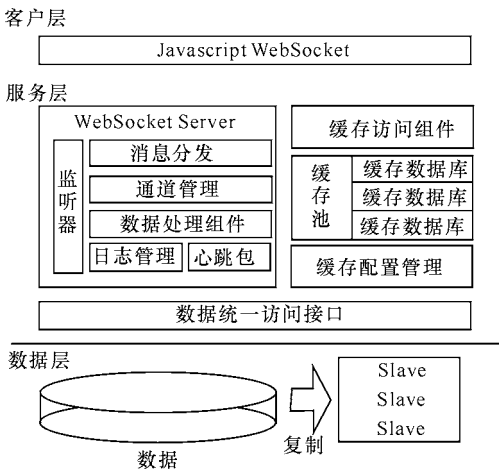


图3 应用框架示意图

### 3 应用实例

广西地震信息共享系统在开发时采用此架构式,Web 层采用 .NET, DB 层采用 SQL SERVER 2005,基于 MS-SQL Server 的 SqlCacheDependency 特性,使得我们可以避免“数据过期”的问题,它能够根据数据库中相应数据的变化通知缓存,并移除那些过期的数据<sup>[3]</sup>。NET 自带的缓存虽然能够给中小型、单一服务器的应用系统带来高效的数据访问性,但对于大数据级别的应用则效率不高,特别是分布式的数据访问,更加束手无策。而 Memcached 是高性能的分布式内存缓存服务器,如图 4 所示,通过把缓存技术与应用系统进行集成,缓存全局对象数据、数据库查询结果等,能够减少数据库访问次数,以提高动态 Web 应用的速度、提高可扩展性,还大大降低了数据库的负载。

地震信息速报发布共享系统的核心业务功能是

地震信息的速报,速报功能分为客户端实时的从服务端得到最新的地震数据,数据库在数据更新后实时将最新数据更新至缓存,在高并发的情况下,用户端实时的请求数据由 Web 缓存提供,这个功能是在 WebSocket 的协议及技术规范中进行开发,具有实时、高效等特点。

采用此架构的 Web 应用系统能够大大的改善数据的查询读取压力。在研究 Web 缓存性能时,一般考虑 2 个评价指标:命中率 HR 和字节命中率 BHR,本研究主要对命中率进行评价。对系统登录分别进行 500、1000、3000 及 4000 次随机登录,实测的缓存命中数分别达到 119、377、1441 及 2328 次,命中率分别达到了 23.8%、37.7%、48% 及 58.2% (图 5),随着浏览数据次数的增加,缓存命中率越来越高。缓存命中率高证明了在读写内存中的数据次数更多,减少了物理读取硬盘的次数。

在实际应用中,缓存式 web 架构能有效并极大地降低数据库的负载,同时又能提高 web 的运行性能。当然这些架构还可以根据具体的应用环境进行变种,以达到不同硬件条件下性能的最优化。

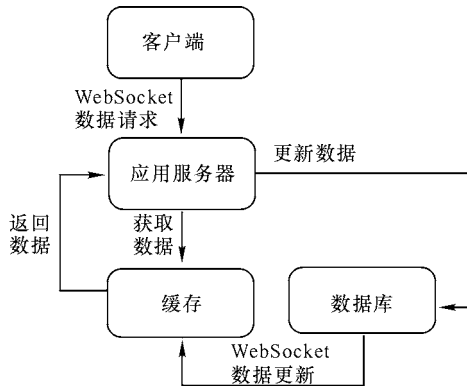


图4 地震信息速报发布共享系统数据请求模型

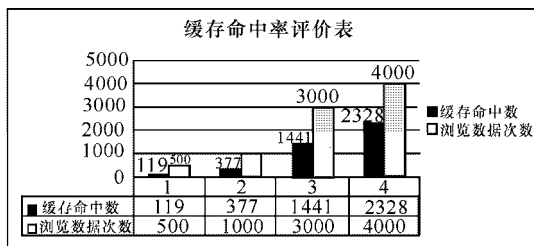


图5 缓存测试示意图

### 4 结束语

运用 WebSocket 的缓存架构能够解决极高并

(下转第 38 页)

资源,满足本层各个引擎的需求。

### 2.3.3 SAAS 建设内容

利用 PAAS 制定企业总线 SOA 服务。用户通过 PAAS 层的定制,能够从 SAAS 层获得与需求相应的服务。服务包括文档上传、安全控制、转换、压缩、版本控制、索引、全文检索、权限管理、用户管理、数据分析和数据维护。这些服务囊括了制造业版式工业文档的一般性需求。用户可以捆绑式获得部分或全部应用服务。

### 2.3.4 客户终端层建设内容

随着移动设备的兴起和硬件能力不断提高,本层不能仅仅保留原有的桌面系统所能利用的功能,客户终端层必须建立包含对移动终端的服务应用支持的接口。

## 3 结束语

本研究以制造业版式工业文档软件功能需求为基础,提出利用服务器虚拟化技术构建云平台的总体架构,为进一步设计和实现制造业版式工业文档管理的云服务平台建设提供了基础。当然这种结构还存在一定的不足,如 IAAS 层的服务还需要针对

具体需求后再调研,PAAS 层还需进一步具体细化建设,由于需求瞬息万变,调整各层次建设的内容时,必须要以客户的实际需求为准进行调整。

### 参考文献:

- [1] Wikipedia. Cloud computing[EB/OL]. [2013-08-10] [http://en.wikipedia.org/wiki/Cloud\\_computing](http://en.wikipedia.org/wiki/Cloud_computing).
- [2] 工业和信息化部电信研究院. 云计算白皮书[EB/OL]. [2013-08-10] <http://www.miit.gov.cn/n11293472/n11293832/n15214847/n15218338/n15224998.files/n15224997.pdf>.
- [3] Yoshihiko Oguchi, Tetsu Yamamoto. Server virtualization technology and its latest trends[J]. FUJITSU Sci Tech J, 2008,44:47-47.
- [4] VMware vSphere. What's new in VMware vSphere 5.5: Platform[EB/OL]. [2013-08-10] <https://www.vmware.com/files/pdf/vsphere/VMware-vSphere-Platform-Whats-New.pdf> (Accessed January, 24, 2014).
- [5] 俞乃博. 云计算 IaaS 服务模式探讨[J]. 电信科学, 2011(S1):98.

(责任编辑:尹 闯)

(上接第 34 页)

发应用对 DB 的访问需求<sup>[4]</sup>;提高系统访问速度,增加系统的 PV 值;利用缓存技术降低对 DB 服务器成本需求,分析业务适用性,哪些业务和场景适合使用,通过缓存技术的研究测试,优化其配置使用,在性能和稳定性上都达到最佳,在现有的功能基础上,自主研发缓存机制统一管理,统一监控,故障排除,实现功能的分布式,开发出一套真正适合我们业务需求的分布式缓存系统,提高现有硬件资源的利用率,发挥硬件应有性能,控制成本增加。但由于 WebSocket 仍然是 HTML 5 中的一个草案,版本变化非常大,它的规范和 API 还是有变动的可能。缓存在框架中的作用受制于缓存策略、脏数据等问题,应结合实际选择更优的技术方案,由于篇幅问题本文

没有一一将技术细节进行挖掘,但可以以为以后 Web 应用系统的构建提供参考。

### 参考文献:

- [1] 何坤. 基于内存数据库的分布式数据库架构[J]. 程序员, 2010(7):116-117.
- [2] 梅宏, 申峻嵘. 软件体系结构研究进展[J]. 软件学报, 2006,17(6):1257-1257.
- [3] 周礼. c#和 .NET 3.0 第一步[M]. 北京:清华大学出版社, 2008.
- [4] 李代立, 陈榕. WebSocket 在 Web 实时通信领域的研究[J]. 电脑知识与技术, 2010(28):7923-7925, 7935.

(责任编辑:陆 雁)