



(12) 发明专利申请

(10) 申请公布号 CN 103457841 A

(43) 申请公布日 2013. 12. 18

(21) 申请号 201310425410. 6

(22) 申请日 2013. 09. 17

(71) 申请人 北京京东尚科信息技术有限公司
地址 100080 北京市海淀区苏州街20号2号楼2层

(72) 发明人 李睿

(74) 专利代理机构 中原信达知识产权代理有限
责任公司 11219
代理人 刘光明 陆锦华

(51) Int. Cl.

H04L 12/58 (2006. 01)

H04L 29/08 (2006. 01)

G06F 17/30 (2006. 01)

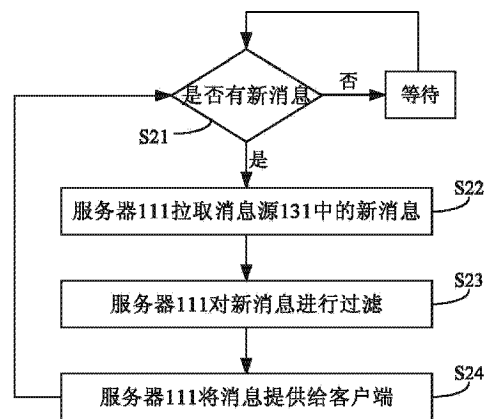
权利要求书2页 说明书5页 附图1页

(54) 发明名称

一种基于长连接的消息处理方法和消息处理装置

(57) 摘要

本发明提供一种基于长连接的消息处理方法和消息处理装置,在基于长连接的消息处理中能够实现跨浏览器平台处理,并且对客户端的容量较大。该消息处理方法包括:对于服务器集群中的各个服务器,该服务器使用一个或多个线程管理与自身与客户端的长连接,每个所述线程用于维持预设数目个所述长连接;所述服务器获取来自于消息源的消息,然后将该消息提供给订阅该消息的客户端。



1. 一种基于长连接的消息处理方法,其特征在于,包括:

对于服务器集群中的各个服务器,该服务器使用一个或多个线程管理与自身与客户端的长连接,每个所述线程用于维持预设数目个所述长连接;

所述服务器获取来自于消息源的消息,然后将该消息提供给订阅该消息的客户端。

2. 根据权利要求1所述的消息处理方法,其特征在于,所述将该消息提供给订阅该消息的客户端的步骤包括:

所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器连接,然后将该消息发送给该客户端。

3. 根据权利要求1所述的消息处理方法,其特征在于,所述将该消息提供给订阅该消息的客户端的步骤包括:

所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器集群中的另一服务器连接,然后将该消息发送给该另一服务器,以供该另一服务器将该消息发送给订阅该消息的客户端。

4. 根据权利要求1,2或3所述的消息处理方法,其特征在于,所述将该消息提供给订阅该消息的客户端的步骤之前,还包括:所述服务器根据所述客户端的配置信息对该客户端订阅的消息进行过滤。

5. 根据权利要求1,2或3所述的消息处理方法,其特征在于,所述将该消息提供给订阅该消息的客户端的步骤之后,还包括:

所述服务器确认该消息发送失败,然后记录发送失败的时间;

所述服务器确认与该客户端的长连接恢复,然后向该客户端发送包含所述发送失败的时间的丢包消息,以供所述客户端根据所述丢包消息查询所述发送失败的消息。

6. 一种基于长连接的消息处理装置,设置在服务器集群中的各个服务器中,其特征在于,所述消息处理装置包括:

长连接管理模块,用于使用一个或多个线程管理所述服务器与客户端的长连接,每个所述线程用于维持预设数目个所述长连接;

获取发送模块,用于获取来自于消息源的消息,然后将该消息提供给订阅该消息的客户端。

7. 根据权利要求6所述的消息处理装置,其特征在于,所述获取发送模块还用于:

从所述服务器连接的消息源获取消息,确认订阅该消息的客户端与所述服务器连接,然后将该消息发送给该客户端。

8. 根据权利要求6所述的消息处理装置,其特征在于,所述获取发送模块还用于:

所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器集群中的另一服务器连接,然后将该消息发送给该另一服务器;以及,

接收另一服务器发送的消息,然后将该消息发送给自身连接的订阅该消息的客户端。

9. 根据权利要求6,7或8所述的消息处理装置,其特征在于,所述获取发送模块还用于对获取的所述客户端订阅的消息根据所述客户端的配置信息对该客户端订阅的消息进行过滤。

10. 根据权利要求6,7或8所述的消息处理装置,其特征在于,还包括失败处理模块,用于在所述获取发送模块将客户端订阅的消息发送给该客户端之后,确认该消息发送失败,

然后记录发送失败的时间,以及确认与该客户端的长连接恢复,然后向该客户端发送包含所述发送失败的时间的丢包消息,以供所述客户端根据所述丢包消息查询所述发送失败的消息。

一种基于长连接的消息处理方法和消息处理装置

技术领域

[0001] 本发明涉及一种基于长连接的消息处理方法和消息处理装置。

背景技术

[0002] 随着互联网技术的发展,各类应用系统中往往需要频繁地将消息发送到客户端,例如在电子商务领域中,很多应用譬如监控、库存变化、即时报价系统都需要将后台发生的变化实时传送到客户端,传统的一种模式是客户发送轮训请求,不停的从服务器获取数据,这种方式同时增加了客户端与服务器的的工作量,因此出现了长连接技术以改进客户端获取消息的方式。长连接推送实现客户端请求服务器,服务器一直保持此连接。服务器从消息源(mq 或者其他消息来源)获取消息,进行分拣(过滤掉一些不重要的字段),最后发送给指定的客户端。相比传统做法,无须客户端不停地刷新、发送请求,这样服务器的压力就会变小,并且可以实时的获取最新数据。再加上使用 nio,并发连接数目很多的场景下对于降低服务器的资源负载非常有效。

[0003] 目前关于长连接推送技术主要都是基于浏览器+ajax 长轮询实现的。一个比较成型的框架是 Pushlet。但是随着系统的不断扩张,终端数目的日益增加、终端所使用的浏览器的种类的不断出新,Pushlet 也暴露出以下不足之处:

[0004] Pushlet 需要使用能工作在任何平台、所有浏览器版本的 DHTML 库。也就是说针对不同的浏览器,需要准备不同的 DHTML 库,存在对浏览器的依赖性;当数目稍大(例如 100 个以上)的客户端通过 Pushlet 连接到服务器时,服务器上的线程和套接字资源都将出现紧张,也就是说。Pushlet 对客户端的容量有限。

发明内容

[0005] 有鉴于此,本发明提供一种基于长连接的消息处理方法和消息处理装置,在基于长连接的消息处理中能够实现跨浏览器平台处理,并且对客户端的容量较大。

[0006] 为实现上述目的,根据本发明的一个方面,提供了一种基于长连接的消息处理方法。

[0007] 本发明的基于长连接的消息处理方法包括:对于服务器集群中的各个服务器,该服务器使用一个或多个线程管理与自身与客户端的长连接,每个所述线程用于维持预设数目个所述长连接;所述服务器获取来自于消息源的消息,然后将该消息提供给订阅该消息的客户端。

[0008] 可选地,所述将该消息提供给订阅该消息的客户端的步骤包括:所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器连接,然后将该消息发送给该客户端。

[0009] 可选地,所述将该消息提供给订阅该消息的客户端的步骤包括:所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器集群中的另一服务器连接,然后将该消息发送给该另一服务器,以供该另一服务器将该消息发送给订阅该消息的

客户端。

[0010] 可选地,所述将该消息提供给订阅该消息的客户端的步骤之前,还包括:所述服务器根据所述客户端的配置信息对该客户端订阅的消息进行过滤。

[0011] 可选地,所述将该消息提供给订阅该消息的客户端的步骤之后,还包括:所述服务器确认该消息发送失败,然后记录发送失败的时间;所述服务器确认与该客户端的长连接恢复,然后向该客户端发送包含所述发送失败的时间的丢包消息,以供所述客户端根据所述丢包消息查询所述发送失败的消息。

[0012] 根据本发明的另一方面,提供了一种基于长连接的消息处理装置。

[0013] 本发明的基于长连接的消息处理装置设置在服务器集群中的各个服务器中,该消息处理装置包括:长连接管理模块,用于使用一个或多个线程管理所述服务器与客户端的长连接,每个所述线程用于维持预设数目个所述长连接;获取发送模块,用于获取来自于消息源的消息,然后将该消息提供给订阅该消息的客户端。

[0014] 可选地,所述获取发送模块还用于:从所述服务器连接的消息源获取消息,确认订阅该消息的客户端与所述服务器连接,然后将该消息发送给该客户端。

[0015] 可选地,所述获取发送模块还用于:所述服务器从自身连接的消息源获取消息,确认订阅该消息的客户端与所述服务器集群中的另一服务器连接,然后将该消息发送给该另一服务器;以及,接收另一服务器发送的消息,然后将该消息发送给自身连接的订阅该消息的客户端。

[0016] 可选地,所述获取发送模块还用于对获取的所述客户端订阅的消息根据所述客户端的配置信息对该客户端订阅的消息进行过滤。

[0017] 可选地,还包括失败处理模块,用于在所述获取发送模块将客户端订阅的消息发送给该客户端之后,确认该消息发送失败,然后记录发送失败的时间,以及确认与该客户端的长连接恢复,然后向该客户端发送包含所述发送失败的时间的丢包消息,以供所述客户端根据所述丢包消息查询所述发送失败的消息。

[0018] 根据本发明的技术方案,采用服务器上运行的程序的线程来维持服务器与客户端的长连接,能够使基于长连接的消息处理不依赖于具体的浏览器,实现了跨浏览器平台处理,因为每个线程能够维持数量级为几千的连接数,因此使服务器能够与大量客户端连接。

附图说明

[0019] 附图用于更好地理解本发明,不构成对本发明的不当限定。其中:

[0020] 图 1 是根据本发明实施例的服务器、客户端以及消息源的连接的示意图;

[0021] 图 2 是根据本发明实施例的基于长连接的消息处理方法的一种流程的示意图;

[0022] 图 3 是根据本发明实施例的基于长连接的消息处理装置的基本结构的示意图。

具体实施方式

[0023] 以下结合附图对本发明的示范性实施例做出说明,其中包括本发明实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本发明的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0024] 图1是根据本发明实施例的服务器、客户端以及消息源的连接的示意图。如图1所示,服务器集群11中有多台互相连接(图中未示出连接关系)的服务器111、112、113、……、11N。其中各个服务器上连接有多个客户端。为示意清晰起见,图中仅示出连接在服务器111上的客户端121和连接在服务器112上的客户端122。实际上各个服务器可以连接有几千个客户端。图1中还示出了多个消息源,同样为了示意清晰起见,图中仅示出连接在服务器111上的一个消息源131和连接在服务器112上的一个消息源132。在实际的系统中,各个服务器可能连接一个或多个消息源,也有可能不与任何消息源连接。

[0025] 对于服务器集群中的各个服务器,该服务器使用一个或多个线程管理与自身与客户端的长连接。服务器可以把多个长连接放入一个池中,并且一个线程管理、维持一个池的长连接。对于现有的处理能力的服务器来说,可以维护多个这样的池,并且每个池可以容纳几千个长连接。另外,各个服务器在运行时,获取来自于消息源的客户端订阅的消息,然后将该消息发送给该客户端。因为服务器采用服务器端的程序线程来实现长连接的维护,这种程序能够独立于浏览器,担当客户端请求与服务器响应的中间层,所以在基于长连接的消息处理中能够实现跨浏览器平台处理。并且这种程序的运行效率主要取决于服务器硬件的处理能力,鉴于目前服务器硬件性能已相当强大,因此运行这种程序时对于对客户端的容量较大,一般每个线程容纳的长连接数目以千计。这种程序可以采用servlet程序。

[0026] 以下结合图2和图3对本实施例中的消息处理方法做出说明。

[0027] 图2是根据本发明实施例的基于长连接的消息处理方法的基本流程的示意图。

[0028] 步骤S21:服务器111在消息源131处查询是否有新消息。服务器111在运行中,如果有需要向客户端发送的消息,则通过自身与客户端之间的长连接进行发送,反之则定期地向客户端发送心跳消息以维持该长连接。如果心跳消息发送失败或者消息发送失败,就确认该长连接中断。该中断可能是由于客户端故障或网络异常导致。在没有需要向客户端发送消息的情况下,服务器111即从消息源处查询是否有新消息。查询结果为有或者没有,从步骤S22起是有新消息的情况。如没有新消息,则服务器111等待一个预设的周期时长之后返回本步骤再次查询,如图中所示。

[0029] 步骤S22:服务器111拉取消息源131中的新消息。

[0030] 步骤S23:服务器111对新消息进行过滤。本步骤为可选步骤,用于客户端的配置信息中有过滤需求的情况。例如客户端要求不接收指定的字段,或者仅接收指定的字段。服务器111相应地从新消息中删除或提取客户端指定的字段。各个服务器都能够获得连接在服务器集群11上的每个客户端的配置信息,不论该客户端是否直接连接在该服务器上。

[0031] 步骤S24:服务器111将消息提供给客户端。这里分两种情况,一种情况是订阅了服务器111所获取的消息的客户端与服务器111直接连接,例如客户端121,这种情况下服务器111根据客户端配置信息确认获取的消息是客户端121所订阅,于是将消息直接推送给客户端121;另一种情况是订阅了服务器111所获取的消息的客户端与另一服务器连接,例如是客户端122订阅了该消息,这种情况下服务器111根据客户端配置信息确认获取的消息是连接在服务器112的客户端122所订阅,于是将该消息推送给服务器112,由服务器112将该消息推送给客户端122。

[0032] 至此,本次的向客户端提供消息的流程结束,服务器111再次在消息源131处查询是否有新消息,即返回步骤S21。

[0033] 如果服务器与客户端之间的长连接中断,则服务器向客户端发送消息时会失败。服务器确认消息发送失败后,丢弃要发送的消息,并且记录消息发送失败的时间。当该长连接恢复并被服务器确认后,服务器向该长连接的客户端发送一个丢包消息,该丢包消息中有记录的消息发送失败的时间,这样客户端就可以根据该时间查询曾经丢失的消息。客户端此时需访问一个网络地址,该地址中包含消息源的消息。并且该网络地址在客户端进行消息订阅时已经获知。

[0034] 图 3 是根据本发明实施例的基于长连接的消息处理装置的基本结构的示意图。以下结合图 3 对本发明实施例中的基于长连接的消息处理装置做出说明。该装置可以设置在图 1 中的各个服务器中。

[0035] 如图 3 所示,消息处理装置 30 主要包括长连接管理模块 31 和获取发送模块 32。长连接管理模块 31 用于使用一个或多个线程管理服务器与客户端的长连接,每个线程用于维持预设数目个长连接;获取发送模块 32 用于获取来自于消息源的客户端订阅的消息,然后将该消息发送给该客户端。

[0036] 获取发送模块 32 还可用于:从服务器连接的消息源获取消息,对于获取的各条消息,确定订阅该消息的客户端,然后将该消息发送给该客户端。

[0037] 获取发送模块 32 还可用于:从服务器集群中的另一服务器获取消息,对于获取的各条消息,确定订阅该消息的客户端,然后将该消息发送给该客户端,其中该另一服务器连接有消息源,该客户端订阅有该消息源的消息。

[0038] 获取发送模块 32 还可用于:对获取的客户端订阅的消息根据客户端的配置信息对该客户端订阅的消息进行过滤。

[0039] 消息处理装置 30 还可以包含失败处理模块(图中未示出),用于在获取发送模块 32 将客户端订阅的消息发送给该客户端之后,确认该消息发送失败,然后记录发送失败的时间,以及确认与该客户端的长连接恢复,然后向该客户端发送包含发送失败的时间的丢包消息,以供该客户端根据该丢包消息查询所述发送失败的消息。

[0040] 根据本发明实施例的技术方案,采用服务器上运行的程序的线程来维持服务器与客户端的长连接,能够使基于长连接的消息处理不依赖于具体的浏览器,实现了跨浏览器平台处理,因为每个线程能够维持数量级为几千的连接数,因此使服务器能够与大量客户端连接。

[0041] 以上结合具体实施例描述了本发明的基本原理,但是,需要指出的是,对本领域的普通技术人员而言,能够理解本发明的方法和设备的全部或者任何步骤或者部件,可以在任何计算装置(包括处理器、存储介质等)或者计算装置的网络中,以硬件、固件、软件或者它们的组合加以实现,这是本领域普通技术人员在阅读了本发明的说明的情况下运用他们的基本编程技能就能实现的。

[0042] 因此,本发明的目的还可以通过在任何计算装置上运行一个程序或者一组程序来实现。所述计算装置可以是公知的通用装置。因此,本发明的目的也可以仅仅通过提供包含实现所述方法或者装置的程序代码的程序产品来实现。也就是说,这样的程序产品也构成本发明,并且存储有这样的程序产品的存储介质也构成本发明。显然,所述存储介质可以是任何公知的存储介质或者将来开发出的任何存储介质。

[0043] 还需要指出的是,在本发明的装置和方法中,显然,各部件或各步骤是可以分解和

/或重新组合的。这些分解和/或重新组合应视为本发明的等效方案。并且,执行上述系列处理的步骤可以自然地按照说明的顺序按时间顺序执行,但是并不需要一定按照时间顺序执行。某些步骤可以并行或彼此独立地执行。

[0044] 上述具体实施方式,并不构成对本发明保护范围的限制。本领域技术人员应该明白的是,取决于设计要求和因素,可以发生各种各样的修改、组合、子组合和替代。任何在本发明的精神和原则之内所作的修改、等同替换和改进等,均应包含在本发明保护范围之内。

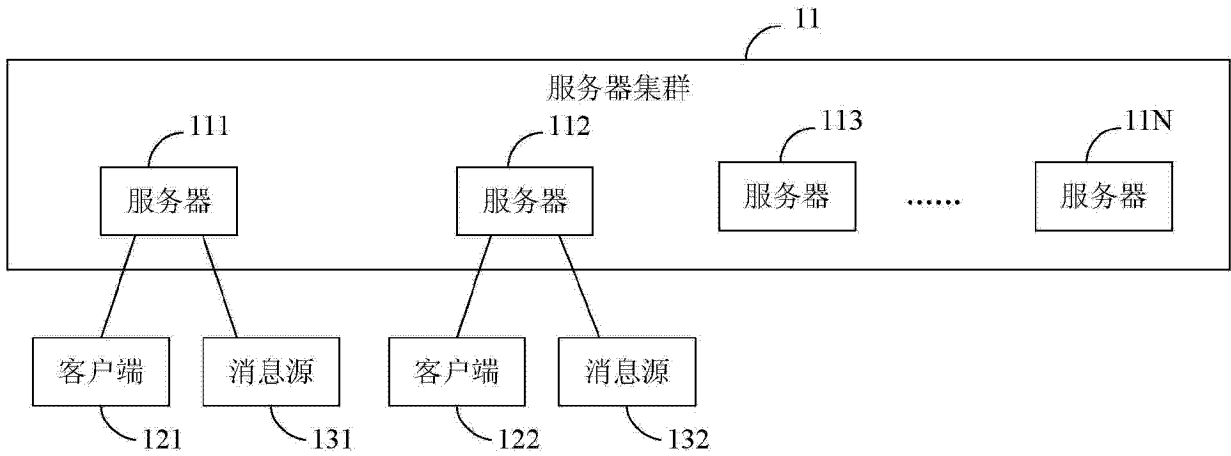


图 1

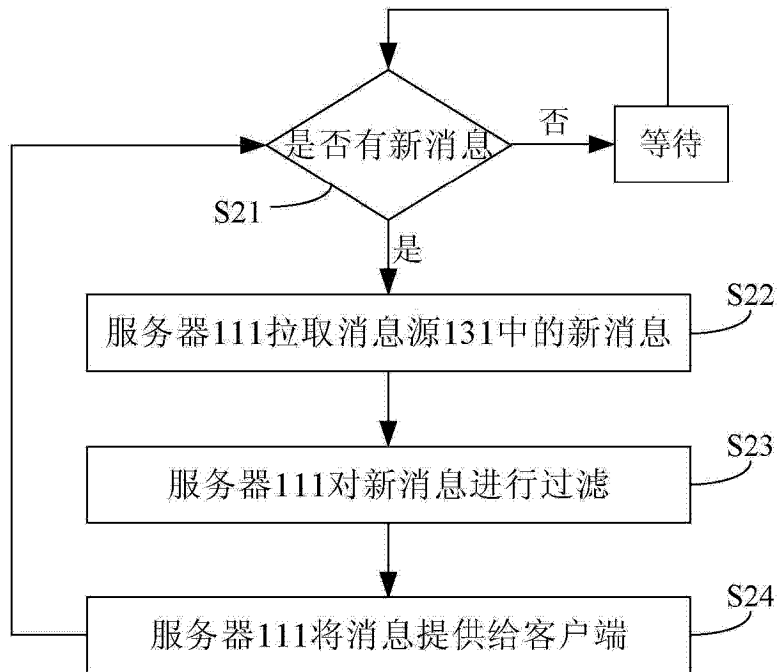


图 2

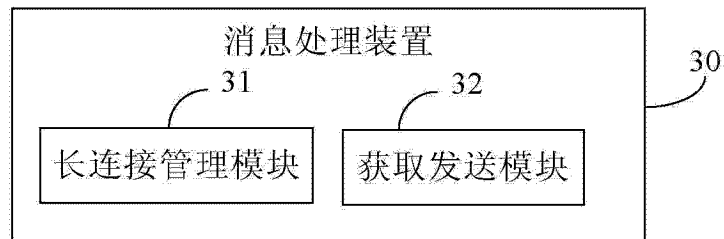


图 3