

一种基于移动代理的服务器集群系统模型

王春枝, 王 骞

(湖北工业大学 计算机学院, 湖北 武汉 430068)

摘 要: 集群是解决网络服务器规模不断扩大的一种有效方案。针对现有的服务器集群的自适应性差、有系统瓶颈等问题, 对采用请求分配器的服务器集群系统进行了研究, 提出一种基于移动代理的服务器集群系统模型, 通过移动代理和层次化网络结构, 降低和分摊通信开销, 并建立集群的基本单位——负载域, 使系统负荷化整为零而又管理有序。并考虑集群系统的特点, 结合蚁群算法改进 Agent 的迁移策略。重点阐述了该模型的设计思路, 并介绍了模型的组成结构和移动代理环境下的负载信息收集、参数的自动调整和网络拓扑结构管理等功能的实现和迁移策略。

关键词: 移动代理; 服务器集群; 负载均衡; 自组织

中图分类号: TP233

文献标识码: A

文章编号: 1673-629X(2009)11-0159-04

A Model of Servers Cluster System Based on Mobile Agent

WANG Chun-zhi, WANG Qian

(School of Computer Science, Hubei University of Technology, Wuhan 430068, China)

Abstract: Cluster is one of the most popular techniques to resolve the problem which the scalability of Internet servers is becoming more and more important. However, most of known Web servers cluster schemes have a lot of disadvantages, such as lack of self-organization, existing bottleneck of system. Proposes a model of servers cluster based on mobile agent, to develop the capability of load-reassignment of servers, and the servers cluster is turned into a hierarchical peer-to-peer network in this scheme. In this scheme, there is a structure called Load Domain, and the whole load on the cluster is distributed to a number of load domains. Introduces the whole idea of design for the cluster, the system architecture and special functions in a mobile agent system.

Key words: mobile agent; servers cluster; load balance; self-organization

0 引 言

随着网络的信息量和访问量成几何级数增长, 服务器超载的问题日益严重, 各类信息中心或数据处理中心对服务器提出了更高的要求, 需要它们既能合理接受各种客户端的请求, 又要能够在很短的时间内做出应答响应。

典型的 Web 服务器集群由一个请求分配器 (dispatcher, 又称前端服务器) 和若干个后台服务器组成^[1]。前端服务器接收外部发来的访问请求, 并通过某种负载分担技术, 将客户请求分配到某一台后台服务器上。这些 Web 服务器集群技术大多采用复杂的算法和信息反馈技术实现负载均衡, 使请求分配器承担了过大的负担, 极易成为性能瓶颈。

文中研究了现有 Web 服务器集群技术, 为进一步解决性能瓶颈问题, 提出了一种基于移动代理的服务器集群方案, 使服务器集群成为一种对等网络, 各台服务器之间可以互相通信, 提高了服务器集群的自适应性和可扩展性。

1 相关研究

移动代理 (Mobile Agent, MA) 是能在同构或异构网络主机之间自主地进行迁移的程序。该程序能自主地决定何时迁往何地, 并可根据需要, 在程序运行的任意一点挂起, 然后迁移到另一台主机上, 并接着这一点继续往下执行。移动代理具有很多属性, 如移动性、自主性、智能性、异构性、主动性、反应性等^[2]。

虽然目前不同移动代理系统的体系结构各不相同, 但几乎所有的移动代理系统都包含移动代理 (简称 MA) 和移动代理服务设施 (简称 MAE) 两个部分。MAE 负责为 MA 建立安全、正确的运行环境, 为 MA 提供最基本的服务 (包括创建、传输、执行), 实施针对具体 MA 的约束机制、容错策略、安全控制和通信机制

收稿日期: 2009-02-18; 修回日期: 2009-05-26

基金项目: 湖北省自然科学基金项目 (2008CDB342, 2008CDZ003);

湖北省教育重点资助项目 (D20081405, D20081402)

作者简介: 王春枝 (1962-), 女, 教授, 硕士生导师, 研究方向为计算机网络与通信、多媒体通信。

等^[3,4]。

移动代理技术给服务器集群系统的设计、实现和维护都带来了新的活力,它有着许多传统集群模型无法比拟的优势。

在传统集群中,前端分配器和后台服务器之间保持着密切的信息交流,有着大量数据在网络中传输,整个系统对某一点的网络带宽存在依赖性,极易形成瓶颈。而移动代理能较大地减轻网络上的数据流量,Agent 携带着任务通过直接移动到目标主机的方式,使得移动代理可以直接访问该主机上的资源,并且减少与源主机的信息交流,规避了通信时延,提高了服务质量。

移动代理具有自主运行能力,能够异步地执行任务。当 Agent 被派遣出去,迁移到目标主机后,就可以断开源主机与目标主机之间的连接,此后移动代理就独立于生成它的进程,并可异步自主操作。当完成相关计算后,再与源主机联系并返回结果。这对今后移动设备或移动服务器来说尤其有用,因为目前移动设备上的计算都依赖于昂贵而脆弱的网络连接,它要求在移动设备和固定网络之间建立持续的连接。无论对于将来的服务器、移动设备还是云计算,移动代理都会是一个有力工具。

移动代理具有较强的应变能力,移动代理能够感知其运行环境,并对环境变化做出适当的反应,它可以根据服务器和网络的负载动态决定移动目标,有利于负载均衡。此外,移动代理的智能路由还减少了用户在浏览或搜寻时所做的判断。

2 基于移动代理的服务器集群模型设计

基于以上研究和分析比较,提出一种基于移动代理的集群系统模型。它能够提供负载均衡的功能,使得多台服务器协同工作,共同完成客户的请求,从而将一组服务器构成一个高性能的、具有高可靠性和良好的可扩展性的虚拟服务器。整个系统的结构对客户透明,无需修改客户端程序。

2.1 设计思想

本集群是在基于负载分配器的集群的基础上,保留分配器的同时引入移动代理技术,因此文中的总体思想有两点:

(1)弱化分配器强化服务器。一方面,分配器采用更为简单的分配算法实现第一次分配,减弱分配器在整个系统中的权重。算法可以采用轮转或随机,这类算法是静态算法,不需要分配器与服务器之间通信,目的是降低分配器的工作负荷和网络带宽的消耗。而保留分配器是为了保持服务器集群对客户的透明性,使

服务器集群有一个统一的对外接口。另一方面,通过引入移动代理,使服务器具有对负载的二次分配能力。单个分配器的能力是一定的,而多个服务器协作所表现出的能力是要强于前者的。当然,协作会带来网络带宽和计算资源的额外开销,所以好的系统结构和算法的设计尤为重要。

(2)服务器分层次。在服务器集群中,服务器的数量可以从几台到几百台甚至上千台,服务器的计算能力和容量也可以是千差万别的,因此,服务器集群不能只存在一个层次,要对服务器进行区别,也就是创建层次^[2]。笔者认为,双层或多层服务器集群结构是有必要的。根据一个统一的标准,每台服务器计算出自己的权值,根据权值,拥有相应的负载职责。而且,围绕某一能力较强的服务器,几台邻近服务器可以组成一个局部负载域,中心服务器要担负一定的控制任务。

根据以上所述,设计如图 1 所示。

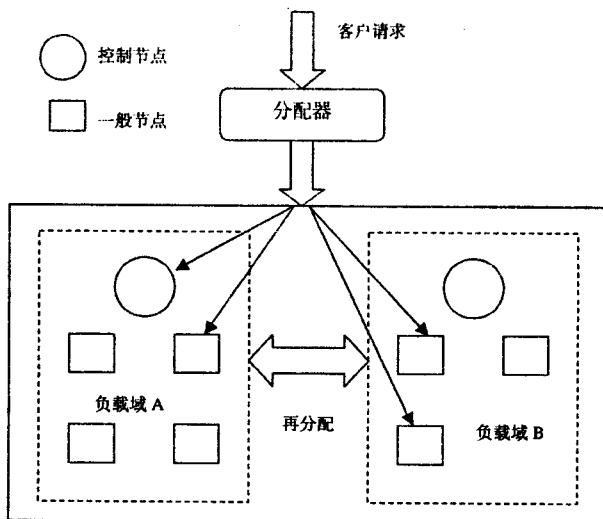


图 1 网络结构图

根据分布性和有序性的思想,设置了负载域和控制节点。邻近节点可以构成负载域,负载均衡优先在负载域内进行,避免了由于 Agent 可能的过度迁移而带来的网络带宽的额外消耗。可以看出,本服务器网络结构不是完全松散的,这和生命体是类似的,一个负载域相当于一个细胞,域内具有一定的自治性,域和域之间有信息交流,同时域还依附于整个系统。负载域内的区域自制和负载均衡的实现是通过控制节点的引入。根据有层次的网络思想,一部分节点具备管理功能,负责区域内的协调调度任务。这一方面是为了更有效的组织服务器,提高系统协同工作的效率;另一方面,是为系统维护人员管理整个负载系统提供方便。事实上,每个安装了集群系统的节点都拥有控制组件,根据网络情况 and 自我状态,可以开启和关闭控制功能。当一个节点开启控制组件的功能后,它就由一般节点

成为控制节点,同时担负响应客户请求和维护负载域的工作。

2.2 系统组成

系统组成介绍(如图 2 所示):

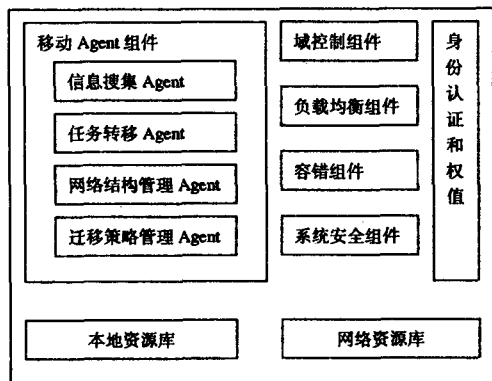


图 2 系统结构图

(1) 负载均衡组件。

负载均衡组件主要用来实现一些底层的负载均衡操作,如进程迁移,包括目标节点的选择、被迁移进程的选择以及迁移的时机等。这些功能的实现需要利用相关的均衡策略以及系统参数。

(2) 移动代理组件。

移动代理组件提供了移动代理的运行环境。同时,用户可以根据需要,通过移动代理生成环境,修改现有的移动代理,或者创建新的移动代理。移动代理组件接收系统主控程序分派的任务,创建并派遣相应的移动代理予以实现,它可以与其余节点交互。

(3) 域控制组件。

域控制组件为控制节点提供了管理负载域的工具,包括负载域内的管理和负载域间的交互。域控制组件的开启和关闭受到身份和权值组件的控制,当节点获得控制权后才开启域控制组件。

(4) 系统安全组件。

移动代理为对等网络提供了更开放的能力的同时,也增加了系统安全的危险隐患。系统安全组件从 3 个方面保护系统:身份认证、权限管理和信息监控。

(5) 容错组件^[5]。

移动代理的迁移主要通过网络数据传输完成,因此,很有可能会发生传输故障,单纯地依赖网络协议进行 Agent 迁移,很难保证其正确性和完整性。因此,必须提供迁移的容错处理服务。

2.3 系统功能

引入移动代理的意义在于它能够实现一些特殊的功能,帮助服务器集群协同工作,包括:异步收集负载信息、自动调整参数、动态调整策略、拓扑结构管理和任务的转移^[6]。

传统方式下,一个节点(如 A)如果要收集多个节点(如 B,C,D)的负载信息,需要分别与各个节点进行通信,通过消息传递,分别获取相关的负载信息。移动代理方式下,一个节点(如 A,即移动代理的源节点)可以发出一个用于收集负载信息的移动代理 MA-X。MA-X 可以自动移动到各个目的节点收集负载信息,最后将结果返回源节点。整个过程中,源节点只需要进行一次发送和一次接收操作。因此,大大减小了源节点的额外负担,提高了效率和性能。

为了保证服务器的服务质量和系统效率,会利用到一些参数,如网络流量、负载阈值、资源利用率边界值等。这些参数的值往往是根据使用经验事先给定的。在负载均衡系统运行的过程中,很多变化无法预料(包括节点状态、系统整体状况、任务执行状况等),因此必须根据需要,及时的对相关参数做出调节,以便从整体上保证负载均衡系统的适应性和性能,也因此提高了系统的应变能力。

针对不同的服务器集群和不同状况下的特殊要求,可以有不同的均衡策略,可以根据需要,从备用策略中,选出一个更加适合的策略,来代替当前的正在服务的策略,该决策由控制节点 n 做出,并派遣策略替换移动代理到各个节点完成该任务(激活或启动备用策略)。当有了新功能时也一样可以通过移动代理去部署服务器。

在该系统中,允许服务器主机动态地加入和退出集群系统。当一个新节点 A 要加入系统时,新节点 A 可以发送一个拓扑移动代理。通过移动代理的迁移,可以发布新节点的信息,如节点号、节点的处理能力、节点的网络地址等。这样该节点就成为了系统的一员,可以共享整个系统的资源,并可与其余节点进行负载均衡了。如果某个节点 B 通过正常的方式退出服务系统,它也可以通过上述类似的方式,来通知各个节点,以避免其余节点将任务转移到该节点。如果节点 B 非正常退出(如死机,断电等),也很容易被发现。因为,会有 Agent 定期的检查各个节点的激活状态,若发现节点 B 在一段时间内没有任何响应,则可以认为该节点非法退出。同样,Agent 把这个消息通知其余节点。

另外,如果节点 A 需要转移某个任务到另外一个节点 B,节点 A 需要和节点 B 协商,在节点 B 同意后,才能转移该任务。如果失败,还需要重新选择其余节点,重复上述操作。系统中,节点 A 只要发送携带了任务信息和若干备选节点信息的 Agent,依次询问目标节点,直到找到合适节点并返回结果。上述过程由移动代理自动完成,大大降低了节点 A 的负担,体现

了移动代理的智能性。

2.4 迁移策略

为了有效地将移动代理应用于集群网络,一个好的迁移策略是个关键因素。使得 Agent 的移动策略在满足网络服务对于集群内信息交流、网络维护和负载均衡的要求的同时,还要尽量少地占用网络带宽和计算资源,真正提高集群的服务质量^[7,8]。

蚁群算法具有较强的稳定性、分散性和鲁棒性,算法中的信息素不断加强,采用了正反馈原理,并利用随机选择策略,加快进化过程。系统结合了服务器集群的特点和蚁群优化算法的优势,研究合适的迁移策略。

考虑三个参数:集群中的网络流量、代理服务信息素浓度和目的主机负载状况。还可以针对不同实时系统的需求综合考虑更多的特征参数。

首先,依次取出节点 Q_i 的邻居节点 Q_j ,计算出 Q_j 关于 Agent A_i 的信息素浓度,然后计算出 A_i 转移到各个服务器节点的概率,再根据流量模型计算出目的服务器节点在未来某时刻的可能网络流量,最后通过概率与流量的比值选出最佳的迁移目的节点。

但是,该策略可能导致移动代理在两个节点之间来回迁移从而使得移动代理无法向着更合适的运行环境前进,所以作了进一步的改进。在计算概率时可引入一个参数 c ,当服务器 k 是上次访问节点时 c 为 0,否则为 1。为了进一步完善该策略,还采用了信息素挥发系数,随着时间的推移对信息素进行调整,防止其无限累加。

3 结束语

研究了现有的基于前端请求分配器的服务器集群系统,为了解决分配器的瓶颈问题和集群的负载均衡问题,提出了基于移动代理的服务器集群系统模型。在该系统中,采用静态调度算法简化请求分配器的工

作内容,引入移动代理技术提高服务器集群的自适应性和负载再分配能力;降低请求分配器成为系统瓶颈的可能性,增强服务器的智能功能:异步收集负载信息、自动调整参数、动态调整策略、拓扑结构管理和任务的转移。使得整个系统在工作强度的分配上更加合理。另外还研究了在集群系统中移动代理的迁移策略,采用蚁群算法,综合考虑了集群中的网络流量、代理服务信息素浓度和目的主机负载状况三个因素,并且设置参数避免了移动代理在两个节点之间来回迁移的问题,还通过信息素挥发系数赋予信息素时间属性,防止其无限累加。

参考文献:

- [1] Chen C, Tsai K. The server reassignment problem for load balancing in structured P2P systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2008, 19(2): 234-246.
- [2] Tsutzbach D, Rejaie R, Sen S. Characterizing unstructured overlay topologies in modern P2P file-sharing systems[J]. IEEE/ACM Transactions on Networking, 2008, 16(2): 267-280.
- [3] 谢毅,周晓峰. 一个基于移动 Agent 的 LBS 系统[J]. 计算机技术与发展, 2006, 16(5): 177-179.
- [4] 吴湘宁,汪渊. 基于群智能的 P2P 计算网格负载均衡算法[J]. 计算机工程, 2007, 33(24): 88-90.
- [5] 王勇,王忠群,韦良芬. 移动 Agent 的一种支持安全与容错的迁移机制[J]. 计算机技术与发展, 2007, 17(3): 169-171.
- [6] 柏海寰,蒋俊杰,汪为农. 基于分散式查找的 Web 服务器集群[J]. 计算机工程, 2006, 32(2): 96-97.
- [7] 林彤,钱华林. 一种节点自组织的多层组播体系结构[J]. 计算机工程, 2008, 34(8): 102-104.
- [8] CardeUini V, Colajanni M, Yu P S. Dynamic Load Balancing on Web Server Systems[J]. IEEE Internet Computing, 2005, 3(3): 28-39.

(上接第 14 页)

- [3] 田宇. 物流服务供应链构建中的供应商选择研究[J]. 系统工程理论与实践, 2003, 23(5): 51-52.
- [4] 王荣培,万麟瑞. 多专家 AHP 的算法改进及其在供应商选择模型中的应用[J]. 计算机应用与软件, 2005, 22(7): 89-90.
- [5] Saaty T L. The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation[M]. New York: McGraw-Hill, 1980.
- [6] 侯书森,孔淑红. 企业供应链管理[M]. 北京:中国广播电视出版社, 2002: 185-188.
- [7] 何炬. 供应链管理中的供应商选择机制[J]. 科学与科学技术管理, 2001(9): 62-65.
- [8] 李博. THPP 决策支持系统的研究与实现[D]. 北京:清华大学, 2006.
- [9] 于江,杨德礼. 供应链管理模式下企业外包设计研究[J]. 大连理工大学学报, 2003, 24(2): 62-66.
- [10] 陈秋荣,周水银. 生产运作管理的理论与实践[M]. 北京:中国人民大学出版社, 2004.
- [11] 邱发智. L 公司外包供应商的评价与选择[D]. 天津:天津大学, 2005.