

基于代理的高性能目录服务的设计和实现^{*}

李 晖¹ 赵曦滨^{2,1} 顾 明¹ 雍建平²

(清华大学软件学院 北京100084)¹ (江苏大学计算机科学与通信工程学院 镇江212013)²

摘 要 目录服务能有效解决虚拟组织中动态多样的资源管理问题。在分布式系统环境下,目录服务的可用性和安全性必须得到保障。本文提出了一种基于代理的目录服务模型,该模型以基于虚拟组织的分布式商业应用为背景,在提高目录服务自身保护能力的同时,保障了目录服务的稳定和高效。

关键词 虚拟组织,目录服务,代理,分布式系统

Proxy-Based Highly Performance Directory Service

LI Hui¹ ZHAO Xi-Bin^{1,2} GU Ming¹ YONG Jian-Ping²

(School of Software, Tsinghua University, Beijing 100084)¹

(School of Computer Science and Telecommunications Engineering, Jiangsu University, Zhenjiang 212013)²

Abstract Directory service can effectively manage dynamic diverse resources in virtual organization. In distributed system, availability and security of directory service are important. The article presents a new proxy-based directory service. The model can support the distributed commercial application based on virtual organization. It strengthens the self-protection ability of directory service while assures the service's stabilization and high performance.

Keywords Virtual organization, LDAP, Directory service, Proxy

1 引言

目前,大规模的社会化生产蓬勃发展,为了提高企业生产效率,增强企业竞争力,企业必须降低产品开发成本,进行重组或合作,集团化已经成为企业发展的必然趋势。在集团化运作环境中,分布在各地的子单位为了共同的生产目标,必须共享企业资源,协同工作。这种集团化企业的信息集成系统具有两个明显特征:基于分布式系统,涵盖更广的地理范围;在此范围内提供海量数据共享。虚拟组织模型的应用为集团化生产中出现的数据共享,协同工作等问题提供了有效的解决途径,加速了现代企业信息系统的运作。

虚拟组织,Virtual Organization(VO)指为了完成大规模生产而协同工作的多个企业或组织所创建的具有共同目标的团体,最初由美国里海大学于1991年在其“21世纪制造企业的战略”的报告中提出^[1,2]。虚拟组织由任务驱动,其子企业、成员不受生产要素所在物理位置和企业组织间的从属关系的约束,利用网络,共享资源,推动任务的完成。它的出现保证了在低成本、高收益下组织的灵活性,为集团化生产和大型项目的顺利完成提供了有效途径。但是,目前VO的应用还面临以下问题:1)VO如何形成,即如何部署资源和服务以形成VO;2)VO如何运作,即为推动某任务的完成如何选择VO中的资源和服务。可以看出,资源和服务是VO应用不可或缺的因素。VO的高效运作归根结底在于VO资源的合理管理。

VO中的资源管理不同于传统组织资源管理。参与虚拟组织的资源种类繁多,随着虚拟组织的运作,大量的资源被动态地加入和删除。多样性和动态性是虚拟组织资源的最主要

特征。

在VO资源管理的众多应用方案中,目录服务更加适合对动态多样资源的管理。目录服务解决了VO的资源管理,因此目录服务对VO来说至关重要。前面提到的资源定位和选择在目录服务中表现为目录信息的选择(Selecting)和配置(Configuration)^[3]。

在商业应用领域中,稳定性和安全性是目录服务性能好坏的主要衡量标准。一方面,当客户请求频繁时,目录服务也能够保证应答数据的准确和应答速度的稳定。另一方面,商业应用不同于科学计算,利用目录管理的商业数据大部分带有机密性,例如财务数据、订单、客户信息等,这些数据一旦被窃取或破坏,后果将不堪设想。因此,就目录服务系统本身而言,应该有自我保护和自我调节功能,防止目录服务器遭受攻击,并能够在受到攻击的目录服务器崩溃后,自动恢复目录数据。

针对以上需求,很多目录服务的研究已展开。其中,安全性研究主要有目录数据的加密,目录访问权限管理。一部分研究以CAS(Community Authorization Service)为基础,通过权限认证防止资源被非法者访问,该解决方案中资源的所有者将资源权限交给团队,由团队管理访问控制权限^[4]。另一些方案在LDAP目录协议的基础上,对目录数据进行了加密。在目录服务的稳定性和高效性方面,Birrell和Kwok-Yan Lam通过相关研究取得了很大进步。Birrell利用数据的归档策略在硬盘上保存当前目录数据的副本,当内存目录数据收到攻击后,利用副本很快能恢复目录数据,重新启动目录服务^[5]。而Kwok-Yan Lam在稳定的基础上改进了目录服务的响应速度,减少了归档策略对目录服务性能的负面影响^[7]。

^{*} 基金项目:国家“863”项目(No. 2003AA148020, No. 2003AA413031)。李 晖 硕士研究生,主要研究方向为目录服务,软件工程;赵曦滨 博士研究生,主要研究方向为计算机系统安全、软件工程;顾 明 副教授,主要研究方向为操作系统、中间件技术、分布式应用系统支撑平台、电子商务;雍建平 讲师,主要研究方向为软件工程与计算机系统安全。

虽然以上两种研究在目录服务的效率和数据安全上做出了较好的改进,但二者仅就一方面问题提出了解决方案,而忽视了另一面。安全性的研究保证了目录服务的安全,却没有考虑目录服务的性能,而目录服务性能的研究保证了良好的服务性能,却忽视了系统本身的安全。

因此,本文在以上两种研究的基础上,提出了一种基于代理的目录服务模型,该模型以分布式商业应用、虚拟组织为背景,充分考虑了目录服务高效稳定和安全的两方面需求,既保证了目录服务的安全,又实现了目录服务的稳定和高效。

2 分布式系统环境下的目录服务

2.1 目录服务

目录是特殊化的数据库,存放资源的基本信息。一个目录可以存贮很多不同种类的信息,并且可以扩展存贮信息的种类。对目录的操作主要包括读目录和写目录,其中读操作远远多于写操作。通过使用目录服务,用户和应用程序可以通过目录认证授权访问这些信息和资源。国际上已经存在几种目录协议标准,其中比较常见的是 X.500,以及从中分离出的轻量级目录访问协议 LDAP(light directory access protocol)。

和虚拟组织相似,网格也使用目录服务进行资源管理。网格是一种分布式计算基础设施,用来实现动态的、跨组织域的资源共享和问题求解,被称为第三代 Internet^[4,5]。为了解决大量资源的管理问题,网格使用了目录服务技术。网格中特殊化的聚集目录服务能够提供资源、服务的专用视图等信息,允许用户和服务通过提交目录服务请求,找到网格中所需的资源。

虚拟组织的目录服务主要在于成员对虚拟组织资源的访问,以及虚拟组织资源的管理方式。虚拟组织机构模型如图1所示,轻量级目录服务通过注册 VO 资源,向数据交换和工作流引擎提供资源的定位和选择,是 VO 运行的核心。

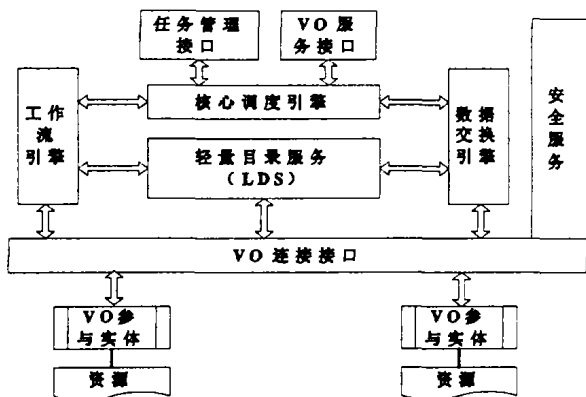


图1 虚拟组织中的目录服务

由于部署过程对系统软硬件环境要求较高,简单地利用 X.500标准或 LDAP 部署实现虚拟组织的目录服务难免显得繁琐。虽然传统的目录服务协议保证了目录的正常运行,但由于在这过程中没有考虑虚拟组织的动态特征,无法使资源管理达到最优。在静态资源和动态资源同时存在的虚拟组织中,为了提供准确、实时的目录服务,目录数据的同步和目录服务的响应速度也受到了广泛的关注。

2.2 相关研究

在前面提到的目录协议标准中,由于 X.500标准的编码机制复杂、对软硬件环境要求太高,从 X.500目录服务标准中分离出来的 LDAP 得到了广泛的应用。LDAP 具有平台无

关、便捷高效等特点。一些研究侧重于如何在虚拟组织中部署和实现 LDAP 目录服务。

对于目录服务安全性的研究,主要有以下两个方面:1)数据加密;2)资源授权。基于 Web 的 LDAP 管理系统受保护的数据分为:传输数据和系统内数据;通过 CA(Certification Authority)和 SSL 传输保证数据的安全^[10]。其他的一些研究集中在 VO 模型的基础上应用 CAS 和 ACL 保证数据安全^[8,9]。数据加密和访问控制是常用的安全手段,前者针对数据本身,后者重点在资源使用过程。

为了在目录数据动态变化的同时保持目录数据的可靠性,Birrell 等人提出了小型数据库(Small Database)的概念^[6]。和商业数据库不同,小型数据库的特点是空间小、数据更新率低。小型数据库应用了归档策略保证目录数据的可靠性。目录服务器中除内存中的目录数据外,还在硬盘上保存内存目录镜像——检查点文件(checkpoint)和记录目录操作的日志文件(log)。当内存的目录数据受损或系统异常被迫重启时,检查点文件和日志文件可以帮助系统恢复内存的目录数据。在小型数据库的基础上,Kwok-Yan Lam 于1997年提出了一种高性能目录服务,对目录服务器的归档过程做了改进^[7]。该方案在一定程度上降低了归档过程对目录响应速度的影响。

以上研究虽然在目录服务的安全和效率方面做了不小的改进,但是安全的目录服务研究只针对目录数据的加密和资源的权限管理,忽略了目录服务系统本身的安全性,也缺少对目录服务稳定性的考虑。而对目录服务系统性能的研究没有加入目录服务系统的安全因素,在这种情况下,高性能的目录服务器一旦受到攻击就会完全瘫痪。

针对目录系统本身,本文提出了一种基于代理的目录服务模型,在对模型改进的过程中,我们发现,该模型不但保证了目录系统的安全性,而且在提高目录服务性能方面取得了较大的进步。

3 基于代理的目录服务

传统的目录服务系统一般由客户端和服务端构成,客户端发送目录服务请求,服务器通过接收、回应该请求,提供目录服务。基于代理的目录服务模型利用“隔离”——目录服务器和客户隔离、“备份”——保存目录数据的多个副本的思想,保证了目录服务系统的安全。并在这种安全保障机制上,使用调度策略提高系统的利用率,最终实现了一个安全、高效的目录服务系统。

3.1 系统架构

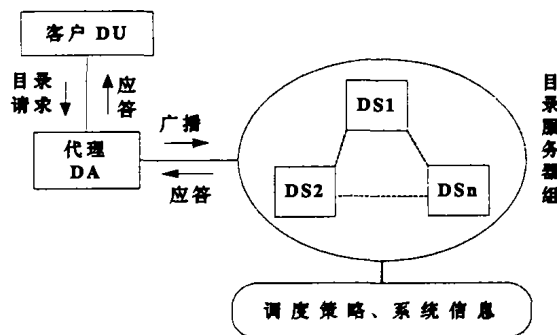


图2 基于代理的目录服务架构

基于代理的目录服务系统构架如图2所示,主要包括三部分:客户端 DU、目录服务器代理 DA、目录服务器组——其中

包含目录服务器 DS。在整个系统中,代理作为目录服务系统面向客户的门户,和客户通信,实现了目录服务器和客户的隔离。目录服务器组在各服务器上保存同一目录数据的多个副本。当有客户请求到来时,代理使用广播方式向各目录服务器组发送目录服务请求,各目录服务器根据客户请求同步地执行目录操作,保证目录数据的一致,并向代理回送请求结果。

3.2 安全保障机制

在本系统中,由于代理仅完成存储转发功能,即使受到了 DoS 攻击无法工作,系统也很容易建立新的代理。而位于局域网内,用于完成主要目录服务工作的目录服务器很难受到 DoS 攻击,保证了目录数据的安全。就算在异常情况下,一台目录服务器由于内部异常或外部攻击而无法工作,其他服务器可以替代它,系统性能并不会受到影响。

3.3 算法描述

系统实现过程中,各目录服务器保存相同的目录数据,在性能和算法上也完全相同。目录服务器的算法基于 Kwok-Yan Lam 模型算法,并在其上稍作改动——归档和响应客户服务都由目录服务器完成,如图3所示。

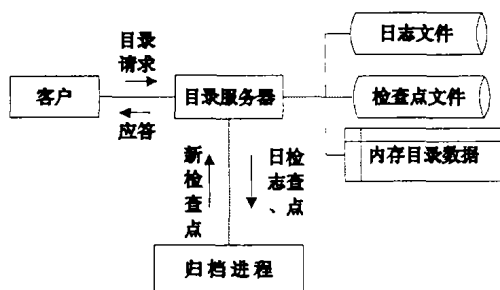


图3 目录服务器的算法描述

目录服务器上保存内存目录、日志文件、检查点文件,其中前者常驻内存,后两者被保存在硬盘上。日志文件记录了每条到来的请求,检查点文件是当前内存目录的副本。为了保证目录服务的稳定和可靠,服务器周期性地归档,更新在硬盘上的检查点文件。

目录服务器上存在两个进程:客户服务进程和归档进程。系统运行时,目录服务器接收客户请求后,启动客户服务进程处理客户请求——执行目录的读写操作,并将结果返回给客户。归档时刻到来时,目录服务器启动归档进程,根据当前日志和检查点文件生成新的检查点文件。一旦内存目录数据被破坏,系统能很快根据检查点文件恢复内存目录数据。

当目录服务器的个数为1时,本模型类似 Kwok-Yan Lam 提出的目录服务模型。当目录服务器个数大于1时,目录服务器组除了实现以上算法,还必须解决两个问题:1)系统负载平衡,也就是如何使系统有较高的利用率;2)目录数据一致性,即分布在各服务器上的目录数据如何保证同步。为了保证各服务器上目录数据的一致,代理使用广播的形式向目录服务器组发送目录服务请求,而系统负载平衡将使用调度策略实现。

3.4 调度策略

良好的调度策略直接影响目录服务的性能。本文调度策略的制定主要考虑以下两方面:首先,该策略能够在保证数据同步的基础上,有效地平衡负载,提高系统利用率;其次,该策略能够保证目录服务的响应速度,也就是实现高效稳定的目录服务。在本模型中有可能影响目录服务性能的因素是归档过程。经过反复的试验比较,我们制定了交替归档调度策略,

在目录响应速度和归档过程之间做了平衡。

交替归档策略如图4所示。首先,根据目录服务器性能 Q_{DS} 和目录大小 D_{DS} , 指定各目录服务器的归档间隔 N_{CK} , 即每收到多少条目录请求归档一次;然后,根据归档间隔 N_{CK} 和系统目录访问频率 F_s , 测试在一次目录数据归档过程中目录服务器得到的目录服务请求数目 N_A ;最后,由归档过程中目录服务的请求个数 N_A 和目录服务器组的服务器个数 N_{DS} , 确定每个服务器的归档时刻,也就是初始化归档计数 N 。使用这种策略,在新的目录服务器加入或旧的目录服务器删除的过程中,修改各服务器归档间隔,对已存在的服务器并无影响。

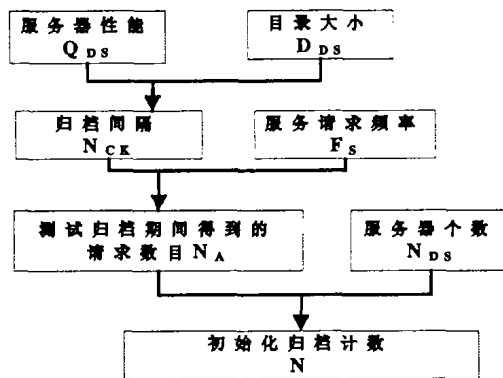


图4 交替归档调度策略

系统运行时,该调度策略的执行如下:

- (1)系统初始化时,就绪的目录服务器加入目录服务器组,并将本机信息(IP地址、开始运行时间等)发送给代理。
- (2)代理接到服务器信息后将其依次存放到内存就绪队列中,并返回应答消息给该服务器。系统初次启动时,代理将首先接受到的目录服务器作为活动服务器。
- (3)各目录服务器根据加入的先后顺序得到不同的初始化归档计数和相同的归档间隔。
- (4)各服务器到达归档时刻后,准备归档并向代理发送归档请求。
- (5)若需要归档的服务器不是活动服务器,代理将其从就绪队列中取出。
- (6)如果活动服务器需要归档时,就绪队列为空,代理不对它的归档请求作任何处理;若就绪队列不为空,代理从内存就绪队列中取出下一个服务器作为新的活动服务器,并向全网广播。
- (7)归档的目录服务器完成归档后,向代理发送就绪请求,代理将该队列重新放回到就绪队列中,等待下次调用。
- (8)当系统信息发生改变时,各目录服务器将修改本地配置文件,代理根据目录服务器信息生成新的内存就绪队列。

举例说明该调度策略的应用。在某个目录服务系统 DSS 中,通过实验测得归档间隔为3000,目录服务器组有两个目录服务器 DS1, DS2, 根据服务器性能、目录大小和目录服务请求频率,最终测试出,归档期间每个服务器大概得到500条目录请求。因此,根据交替归档策略,先加入的目录服务器1的归档初始化计数设为2000,后加入的目录服务器2设为0。

系统运行后,DS1被指定为活动服务器,在收到第1000条目录服务请求时,它的目录计数为3000(3000%3000=0), DS1开始归档。代理从就绪队列中取出 DS2, 将其作为活动服务器,此时 DS2接收的目录请求也为1000条,目录计数为1000。当 DS2接受到第3000条记录时,目录计数为3000(3000%3000=0), DS2开始归档。由于在归档期间得到目录

请求约为500条,因此 DS1此时已经归档完毕被放入就绪队列,代理从就绪队列中取出 DS1作为活动服务器。

上述调度策略中,调度策略和系统信息的控制可以由代理或目录服务器组的配置文件完成。为了减轻代理负荷,增强系统的安全性,使用了目录服务器的本地配置文件保存调度策略和系统信息。

4 试验分析

在理论分析的基础上对本文的目录服务模型的安全性和高效稳定性进行了测试,并通过和 Kwok-Yan Lam 目录服务模型比较,证实了本文的目录服务模型能有效地实现安全、高效的目录服务。

在三台 PC 机上模拟运行了目录服务。其中,一台服务器上部署了代理和客户程序,目录服务器组由其他两台机器组成,分别为服务器1和服务器2,位于同一局域网内。操作系统是用 Windows Advance Server 2000,内存为256M。实验过程中分别在目录记录为10,000、100,000、1000,000、10,000,000条的情况下,测试了目录服务性能。我们以10,000,000条目录记录,目录大小为7.62M的情况为例详细说明实验过程。

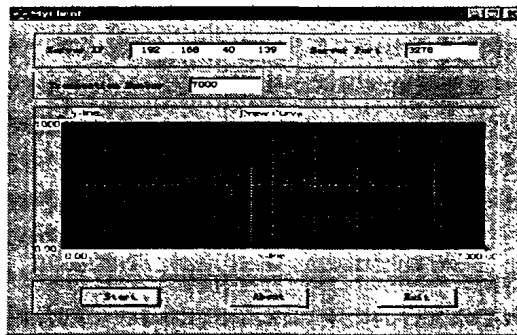
在安全性能测试中,我们首先模拟 DoS 攻击,使代理无法工作。这样,在一定时间内服务器无法接收到代理的任何消息,客户由于接受不到目录请求应答,停止向代理发送服务请求。这时,系统能很方便地配置新的代理,重新开始工作,对目录服务器并不产生任何影响。其次,我们假设活动服务器产生了异常,无法继续服务。这时,代理收到服务器的异常消息后,从就绪队列头中取出另一台服务器继续工作,并不影响目录服务和目录数据的保存,目录服务的性能并没有受到影响。

最后,我们对本文的目录服务模型、Kwok-Yan Lam 模型进行了性能测试和比较。

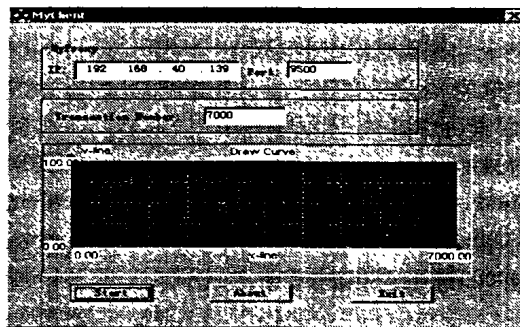
Kwok-Yan Lam 算法中,目录数据的归档由归档服务器完成。当目录服务器上的归档进程启动时,目录服务器向归档服务器传送当前的检查点文件(checkpoint_x.dat, log_x.dat),同时生成新日志文件,继续处理客户请求。归档服务器将归档后的数据——新的检查点文件(checkpoint_x+1.dat)回传给目录服务器。目录服务器接收到数据后,在服务器中生成相应的检查点文件,同时继续处理客户请求。服务器在归档过程中并不会停止对客户请求的处理,但是归档数据的传送和接收仍然影响了对客户请求的响应速度,如图5(a)。

对本文模型测试时,首先在目录服务器上生成目录数据文件和日志文件,以及版本号文件。目录数据文件即为初始化的检查点文件(checkpoint_0.dat),日志文件(log_0.dat)初始化为空,版本号文件(version.dat, newversion.dat)中记录当前日志文件和检查点文件的序列号,初始化时记录都为0。然后,使用交替归档的调度策略,我们确定了目录服务器1的初始化归档计数为2000,目录服务器2的初始化归档计数为0。试验过程中,客户机模拟实际客户,向代理发送请求,目录服务器组根据调度策略和归档策略响应代理发送的请求,并周期性归档。

Kwok-Yan Lam 模型和本文目录服务模型比较如图5。在两种模型的客户端界面上,中央性能图(黑色框)的横坐标表示发送的记录编号,纵坐标表示响应时间。当客户机一共发送7000条请求,服务器每3000条归档一次的情况下,图5(a)记录了 Kwok-Yan Lam 模型的响应速度。图5(b)记录了相同试验环境中,本文的目录服务模型的响应速度。



(a) Kwok-Yan Lam 目录服务模型



(b) 基于代理的目录服务模型

图5 Kwok-Yan Lam 方案和本文方案的比较

可以看出在 Kwok-Yan Lam 方案中,当目录服务器启动归档进程时,会影响当前的客户请求服务速度。大约第3000条记录的响应时间最长,接近750ms。发送7000条记录时,服务器将进行两次归档,图中出现了两个峰值。对 Kwok-Yan Lam 模型反复测试的结果表明,目录数据量越大,归档过程对客户请求响应速度的影响越大。由服务器完成归档过程比由归档服务器完成要高效得多。也就是说,对于大批量数据,网络的传送速度未必比文件的读写和更新快多少。而且,根据目录访问的特点——读操作远远多于写操作,归档的执行——检查点文件的更新也比较省时。

使用本文的目录服务模型,由于两台服务器交替归档,客户将不会明显感觉到服务器归档过程对响应速度的影响,如图5(b)。只有在代理执行活动服务器切换时,对客户的响应速度会有轻微增加,约10ms,是平常响应速度的2~3倍。

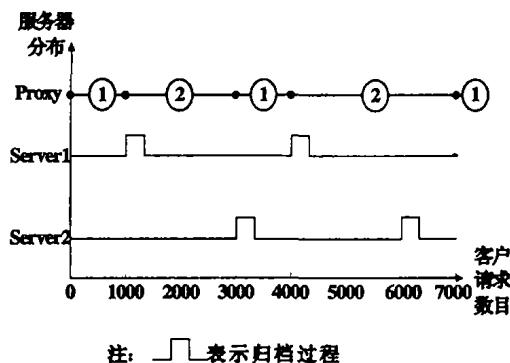


图6 代理和服务器的运行过程

对图5(b)的试验过程,我们在各服务器上进行了跟踪,如图6所示。横坐标为客户请求条数,纵坐标为各服务器的分布。在代理对应的线段上,圆圈表示了某活动服务器,活动服务器的序号已在圈内标明。各服务器的线段区别了服务器的归档时刻和非归档时刻,归档时刻用折线表示。例如,服务器1在1000—1400条请求到来期间归档。

根据跟踪信息,我们清楚地发现各服务器的执行过程。代理服务器将在1000,3000,4000,6000的地方进行服务器切换,如图5(b),可以发现1000,3000,4000,6000条记录附近出现了轻微的峰值,这正是代理在进行服务器切换。例如0到1000期间,服务器1为活动服务器,第1000条请求到来时,活动服务器被切换到服务器2,依次类推。同时,各服务器的归档过程也在响应时刻按时执行。总体来看,这种无缝的服务器切换过程得益于良好的调度策略。

实验结果表明,本文的目录服务模型对目录服务请求的平均响应速度远远小于 Kwok-Yan Lam 模型的平均响应速度。切换服务器的代价远远小于归档带来的影响,这一点可以比较图5(a)、(b)得出。在使用良好的调度策略后,切换服务器所带来的目录响应延迟几乎可以忽略不计。

我们在测试过程中发现,目录的大小和归档线程的间隔是有密切联系的。如果目录较大,归档间隔小,则由于前一次归档未完后继的大部分归档得不到执行。因此,制定恰当的归档策略很重要。同时,该试验结果证明我们的调度策略是非常成功的,根据调度策略制定的初始化归档计数,有效地保证了在某一目录服务环境——目录数据、请求频率、服务器性能等下,交替归档的成功实现。总之,对于相同的目录服务环境,基于代理的目录服务模型不但保证了目录服务的安全,也在一定程度上提高了目录服务的性能,减少了归档过程对目录服务性能的影响。

小结 本文在分析虚拟组织资源管理的动态性、多样性等特点的基础上,结合其资源管理需求,提出了一种基于代理的目录服务模型。该模型保证了目录服务系统本身的安全性,同时在多目录服务器的环境下,利用调度策略有效平衡系统负载,实现了高效稳定的目录服务。在虚拟组织规模较大、资源访问过于频繁的情况下,多台服务器交互保证了目录服务

(上接第68页)

基于 LVQ 网络反垃圾邮件过滤器发现:

1. 网络的训练次数对于 LVQ 的性能有很大的影响,如果没有充分的训练,性能还达不到要求,在训练1500次左右,性能基本达到稳定。

2. 基于神经网络的方法普遍好于贝叶斯算法。分析大概是因为神经网络更加整体考虑了各个特征单词之间的关系,而在贝叶斯算法中,它是过于简单地假设各个特征单词之间是独立的。

3. 基于 LVQ 算法的过滤器优于基于 BP 算法的,是因为我们细化了垃圾邮件的类型,每个类型的垃圾邮件有更加明确的特征向量范围,更利于计算机的识别。

参考文献

- 1 上海艾瑞市场咨询公司. 2004年中国反垃圾邮件研究报告, 2004, 3
- 2 Cohen W W. Learning rules that classify e-mail. In: proc. of the 1996 AAAI Spring symposium in information access, 1996
- 3 Sahami M, Dumais S, et al. A Bayesian Approach to Filtering Junk E-Mail. Learning for Text Categorization -Papers from the AAAI Workshop, Madison Wisconsin. 1998
- 4 Androustopoulos I, Koutsias J, et al. An experimental comparison

的高响应性能,提高了虚拟组织的资源访问速度。

参考文献

- 1 Katzy B R. Design and Implementation of Virtual Organization, System Sciences. In: Proc. of the Thirty-First Hawaii Intl. Conf. on, 1998, 4: 142~151
- 2 Chrysanthis P K, Znati T, Banerjee S, Chang Shi-Kuo. Establishing virtual enterprises by means of mobile agents, Research Issues on Data Engineering: Information Technology for Virtual Enterprises. In: RIDE-VE '99. Proc, Ninth Intl. Workshop on, March 1999. 116~123
- 3 Fitzgerald S, Foster I, Kesselman C, et al. A Directory Service for Configuring High-Performance Distributed Computations, High Performance Distributed Computing. In: Proc. The Sixth IEEE Intl. Symposium on, 1997. 365 ~ 375
- 4 Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid ---- Enabling Scalable Virtual Organizations, Cluster Computing and the Grid. In: Proc. First IEEE/ACM Intl. Symposium on, 2001. 6 ~ 7
- 5 Czajkowski K, Fitzgerald S, Foster I, Kesselman C. Grid Information Services for Distributed Resource Sharing, High Performance Distributed Computing. In: Proc. 10th IEEE Intl. Symposium on, 2001. 181 ~ 194
- 6 Birrell A D, Jones M B, Wobber E P. A Simple and Efficient Implementation for Small Database. In: Proc. of the 11th ACM Symposium on Operating Systems Principles, 1987. 149~154
- 7 Lam K-Y, Salkield T. Implementing a Highly Available Network Directory Service. Systems Software, 1997, 37: 41~47
- 8 Cannon S, Chan S, Olson D, Tull C. Using CAS to Manage Role-Based VO Sub-Groups. In: Computing in High Energy and Nuclear Physics Conf. San Diego, 2003
- 9 Pearlman L, Welch V, Foster I, Kesselman C, Tuecke S. A Community Authorization Service for Group Collaboration. Policies for Distributed Systems and Networks. In: Proc. Third Intl. Workshop on, 2002. 50~59
- 10 Yang C S, Liu C Y, Chen J H, Sung C Y. Design and Implementation of Secure Web-based LDAP Management System, Information Networking. In: Proc. 15th Intl. Conf. on, 2001. 259~264

of Naive Bayesian and Keyword-based anti-spam filtering with encrypted personal messages. In: Proc. of the 23rd annual intl. ACM SIGIR conf. on research and development in information retrieval, Athens, Greece

- 5 Carreras X, Mrquez L. Boosting trees for anti-spam email filtering. In: Proc. of RANLP-01, 11th Intl. Conf. on Recent Advances in Natural Language Processing, Tzigov Chark, BG, 2001
- 6 Chen Duhong, Tong jie, et al. Spam Email Filter Using Naive Bayesian, Decision Tree, Neural Network and AdaBoost. <http://www.cs.iastate.edu/~tongjie/spamfilter/paper.pdf>
- 7 Clark J, Koprinska I, Poon J. A neural network based approach to automated e-mail classification. In: Proc. of the IEEE/WIC intl. conf. on web intelligence
- 8 Salton G, Wong A, Yang C S. A vector space model for automatic indexing. Communications of the ACM, 1975
- 9 Salton G. Introduction to modern information retrieval. New York. McGraw-Hill Book company, 1983
- 10 Church K W, Hanks T. Word association norms, mutual information and lexicography. In: Proc. of ACL27, Wancouver, Canada, 1989
- 11 Hagan M T, Demuth H B, Beale N H. Neural network design. China Machine Press, 2002, 8