

一种基于 MDS4 的资源访问解决方案

曾永健

(华南师范大学 计算机学院, 广东 广州 510500)

摘要:通过对网络流量和应用系统的分析,对于资源的访问占用了绝大部分的网络带宽和消耗了大部分的服务器资源。为了更高效地提高资源访问效率,在分析和研究网络监控和发现组件及其面向服务的体系结构的基础上给出了一种基于 MDS4 的能有效地提高资源的访问效率的解决方案,并阐述了原型系统中主要的实现细节。所提出的方案,对已有系统修改较少,基本不增加原有系统负担,动态扩展性好,不影响原有系统的正常运行,能线性地提高并发度和系统的吞吐量。

关键词:MDS4; 网格计算; Web services; 网格监控; 同步

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2010)07-0048-04

An MDS4 Based Assessing Resource Solution

ZENG Yong-jian

(School of Computer Science, South China Normal University, Guangzhou 510500, China)

Abstract: By the analysis of the flow network and application system, assessing the resource costs most of the bandwidth and the server resource. In order to improve the efficiency of assessing resource, presents an MDS4 based solution on the analysis and research of monitoring and discovery component and service-oriented architecture, and also describes the implementation details of the key part in the prototype system. The solution presented in the paper makes less modification to the old system and doesn't add to the old system's burden. It also has the advantage like good expansibility, less influence on the old system, improves concurrency granularity and throughput of system linearly.

Key words: MDS4; grid computation; Web services; grid monitoring; synchronization

0 引言

随着网络的发展和人们求知的渴望,大用户高并发访问问题凸显,传统的 CS 结构不堪重负,系统反应迟钝。为提高访问的并发度,提高访问速度,人们对其提出了许多解决方案。文献[1]介绍了在网络边缘建立一个完善全面的中间 CDN(Content Delivery Network),源服务器的内容被分发到各边缘缓存模块 Edge Cache 中的内容加速服务器,用户对内容的访问被重定向到离该访问请求最近的内容加速服务器。文献[2]提出了采用“推拉”结合的内容分发方式可进一步减少用户访问的平均等待时间。文献[3]提出了根据用户访问地所在网络, DNS 策略解析(智能 DNS)能把网站的域名分别解析成不同的地址,达到对用户进行分流的目的。文献[4]介绍了利用集群技术来克服单点故障和利用负载均衡器均衡算法对用户进行分

流,提高系统的并发度。上述解决方案在静态数据的访问上取得了一定的成效,但是对于动态数据,如资源索引、查找资源还是需要到数据中心去查找,随着用户数的增加,数据中心成为系统的瓶颈。另外,网络带宽又是一大瓶颈,特别是对于集群技术,当瓶颈出现在源服务器一侧时,集群技术将受限于出口带宽,无论负载均衡器运用哪种调度算法,都收效甚微。文中在分析 Globus 工具集^[5](Globus Toolkit, GT)的监控和发现系统(monitoring & discovery system, MDS4^[6,7])机理的基础上,设计了一种基于 MDS4 的网格资源访问解决方案,该方案基于分布式的环境,能对访问用户进行有效的分流,相当于增加了有效带宽。

1 相关技术

1.1 MDS4

GT 提供了一个完整的开放网格服务基础设施实现。作为 GT 的重要组成部分的监控和发现组件(MDS4),该组件实现了基于 LDAP^[8]的树状元数据目录服务,它简化了分布式环境下的监控和发现资源的

收稿日期:2009-11-05;修回日期:2010-01-30

基金项目:广东省科技攻关项目(2006B15001004)

作者简介:曾永健(1982-),男,福建龙岩人,硕士研究生,研究领域为网格计算。

任务。

理解 MDS4 的关键在于聚合信息资源框架。其基本的思想如下^[5]:

(1)目标资源信息需要向聚合服务注册。

(2)注册时需提供生存周期:如果超过时间没有更新,则资源失效。因此,聚合服务需要提供自清理功能。

(3)聚合器通过信息源提供的访问方式,定期收集所注册资源的最新状态。

(4)聚合器通过 Web Services 接口发布所获得的注册资源的状态信息。

1.2 面向服务的体系架构

面向服务的体系结构(SOA)是一种用于构建分布式系统的方法。SOA 的特点是:(1)松耦合,(2)明确定义的接口,(3)位置透明,(4)无状态服务^[9]。

在基于 SOA 架构的典型应用场景中,服务提供者首先给出该服务的具体实现,用 WSDL 描述,并将其发布到服务注册中心。该服务接口描述可以被遵从 UDDI 规范的服务消费者通过发现操作将其发现。此后,服务消费者发送发现请求就可以从注册中心获得关于服务的描述,通过 SOAP 协议与服务实现之间进行消息交互。图 1 表示服务消费者和服务提供者之间的通讯过程^[10]。

1.3 Web 服务

Web 服务^[10]是由 URI(uniform resource identifier)标识的软件应用程序,其接口和绑定可以通过 XML 构件进行定义、描述和发现,Web 服务支持通过基于 Internet 的协议使用基于 XML 消息与其他应用软件直接进行交互。

Web 服务的关键在于使用基于 XML 的描述机制进行描述,只需要提供该应用程序接口的 WSDL 描述,就可以成为完全符合定义的 Web 服务。

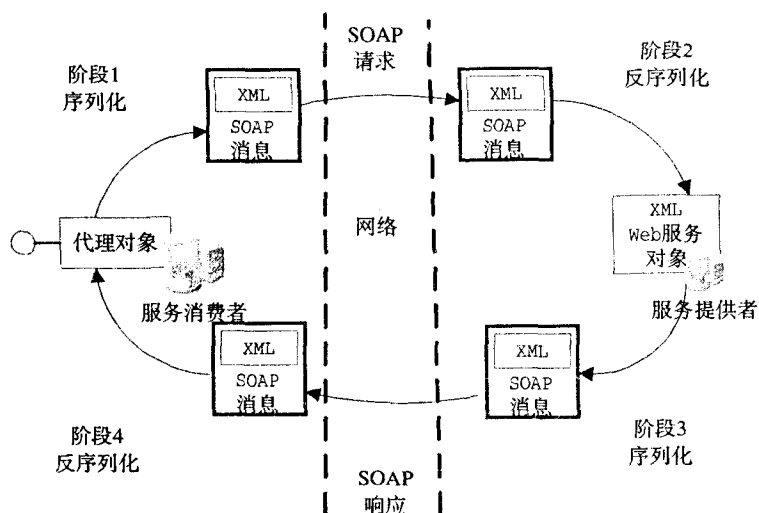


图1 服务消费者和服务提供者之间的通讯过程

2 设计方案

2.1 一个简单的资源访问流程

为了更好的理解资源访问,首先介绍一个基本的资源访问流程图^[11]。如图2,其中MDS服务的作用是收集和发布系统状态信息,NWS服务(Network Weather Service)周期性地监视、动态地预测各种网络性能和计算资源,并且该服务可以在给定时间内把这些信息发送出去。网络资源中的信息采集模块MDS先收集必要信息,然后对资源按照一定的策略进行分发。在调度模块中,根据主机节点的状态,由调度算法计算出最佳的主机来给应用程序提供服务。

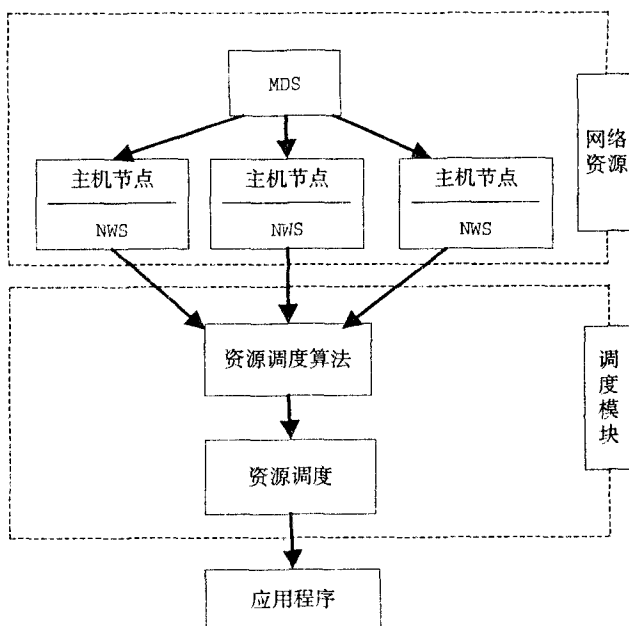


图2 一个简单的资源访问流程图

2.2 基于 MDS4 的网格资源访问设计方案

基于 MDS4 的资源访问解决方案的设计目的是为了更有效地访问资源。具体地说,其原型系统将主要涉及三个方面:网格服务资源的监控、监控数据的分发、基于监控数据的访问决策。设计涉及到以下问题:

(1)如何标识网格节点。网格节点在网格环境中应该唯一标识,我们采用全局统一标识符(GUID)来标识一个节点。GUID是指在一台机器上生成的数字,它保证对在同一时空中的所有机器都是唯一的。

(2)如何获取资源索引变动。资源索引变动,可以通过通知(notify)方式,还可以通过监控方式。通知方式比较及时有效,但考虑到需要对原有系统进行修改,为了更具通用性,这里采用定期监控方式来获取资源索引变化记录。

(3) 如何获取网格节点参数。网格节点数据分为静态数据和动态数据。静态数据例如网格节点描述、位置、网格节点配置等等。动态数据例如内存使用率、CPU 使用率、当前的连接数等等。静态数据可以通过注册方式,当节点加入网格环境的时候,调用中心端的注册服务,利用注册服务直接写入静态信息。动态数据可以通过监控中心周期性地询问各个子网格节点来获得动态数据,也可以通过各个子网格节点定期向监控中心进行汇报。

(4) 以何种形式记录监控数据。因为记录的内容是网格节点的状态信息,信息增删、更新频繁,同时可能需要查阅某些服务器运行状态记录,所以数据库更加适合管理这样的数据。

(5) 信息分发方式。对于需要同步的信息,根据其紧急等级程度不同,分为需要及时通知的和不需要及时通知的。对于一些异常信息,例如:网格节点状态异常、网络通讯中断等信息需要及时通知相关网格节点。对于不需要及时通知的信息,可以加入等待队列中。

我们的设计的方案是,通过中心机房数据库资源索引下发到各校级数据库服务器,监控服务器周期性地获取各个资源服务器的运行状态并把该状态发到各校级服务器节点,用户在校级服务器节点上访问资源时,通过本地查询资源索引表获得资源的相对路径,根据当前资源服务器运行状态信息,自动选择最佳的服务器,拼接上资源索引的相对路径,拼接成最终的访问路径进行访问。方案如图 3 所示。

3 实现细节

为了方便方案的讨论,引入以下定义:

定义 1 作业:指网格环境下,完成一项基本的任务。

定义 2 子作业:隶属于某个作业。一个作业由若干子作业构成。

定义 3 作业状态:指作业当前处于未处理、处理中、已完成、异常等状态。

当子作业中都处于未处理,作业状态为未处理;当

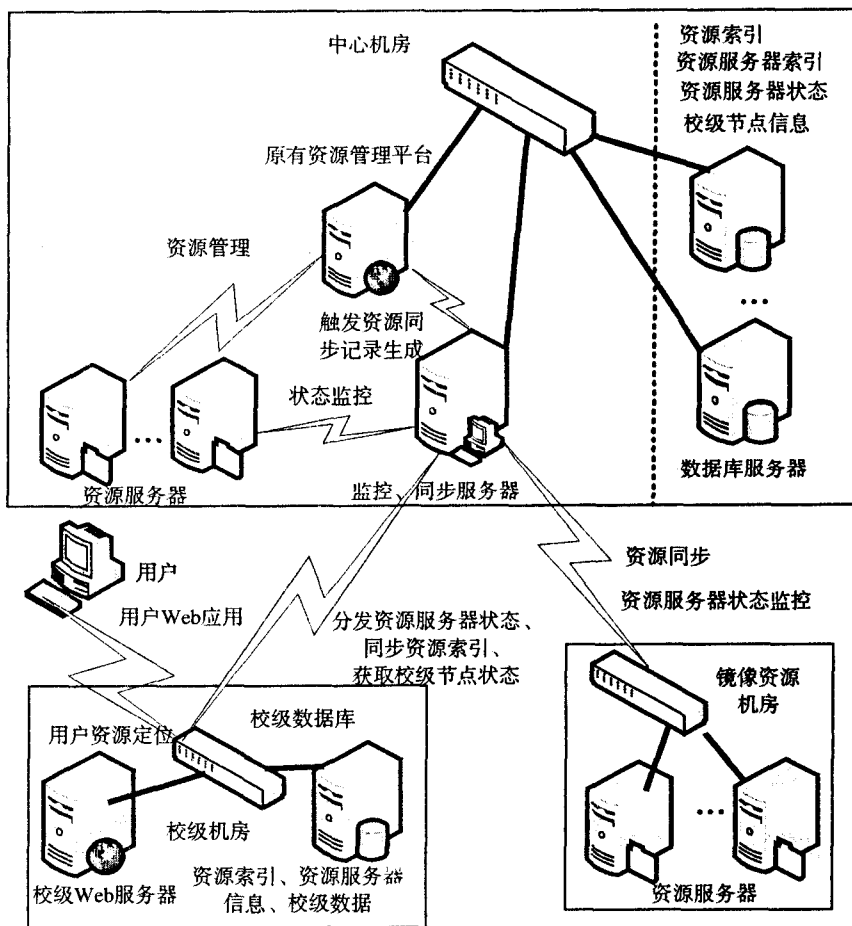


图 3 基于 MDS4 的资源访问解决方案

子作业中存在一个状态为处理中,作业状态为处理中;当全部子作业状态为已完成,作业状态为已完成;当子作业中存在一项状态为异常,作业状态为异常。

定义 4 周期任务:每隔一指定时间运行一次作业调度。

3.1 搭建网格硬件环境

中心机房:部署应用服务器、数据库服务器、资源服务器、监控和同步服务器。

镜像机房:部署镜像资源服务器。

校级机房:部署应用服务器、数据库服务器、校级资源服务器。

3.2 网格环境所部署的 Web 服务

(1) 获取资源索引变动服务:该服务根据传入的日期参数,获得该日期资源索引变动记录。中心机房的数据库服务器发布此服务。

该 Web 服务方法形式: `public ResourceIndex [] getYesterDayIndexChange(Date date);`

(2) 获取网格节点运行状态服务:该服务返回该节点的内存使用率、CPU 使用率、连接数等运行参数。中心机房的资源服务器和镜像机房的资源镜像服务器发布此项服务。

获取内存使用率的 Web 服务方法形式: public float getMemoryInfo();

获取 CPU 使用率的 Web 服务方法形式: public float getCpuInfo();

(3)更新服务:该服务根据传入的对象参数的对象类型,更新对象信息。在校级服务器上部署该服务,实现对校级节点数据的更新。

更新对象信息通用的接口定义如下:

```
public interface UpdateService<T>
{
    public boolean insert(ArrayList arrayList);
    public boolean update(ArrayList arrayList);
    .....
}
```

其中 T 是泛型参数。当 T 是资源服务器类型时,表示该 Web 服务准备添加一系列的资源服务器,或者更新相关的资源服务器信息;当 T 是服务器状态类型时,表示该 Web 服务准备更新或添加相关服务器的状态信息。

3.3 定时任务

文中所涉及的定时任务,包括每天运行一次获取索引变动服务,每 30 分钟运行一次获取各网格节点状态信息服务,每 5 分钟进行一次作业调度服务。

3.4 主要功能模块工作流程

3.4.1 网格节点维护

网格把分散的服务器节点纳入网格环境管理,服务器之间分工和协调工作。

(1)添加校级节点服务器,把各个校级服务器纳入网格管理。

(2)添加中心机房资源服务器和中心机房资源服务器。

添加中心机房资源服务器信息后,在同步服务器的作业队列队尾添加作业,内容为资源服务器信息更新,目的是使各级校级节点更新资源服务器信息变动情况。

(3)添加镜像机房资源镜像服务器。

添加镜像机房资源镜像服务器信息后,在作业队列队尾添加作业,内容为资源镜像服务器信息更新,目的是使各级校级节点更新资源服务器信息变动情况。

3.4.2 资源变动维护

监控服务器每天 0:00 定时到数据库服务器上去提取资源索引变动的记录,在同步服务器的作业队列队尾添加作业,作业内容是资源索引变动。该作业由两个子作业构成,它们是把这些索引所对应的资源文件拷贝一份到各个资源镜像服务器上,和资源文件全

部成功完成后,再把该索引记录本身同步到各校级机房的数据库服务器上。

3.4.3 网格节点状态维护

监控服务器每 30 分钟调用中心机房资源服务器、资源镜像机房资源服务器的 Web 服务来获取网格节点运行状态信息。同时对比服务器当前状态与上个周期的服务器状态,当发现服务器状态发生突变时,比如状态由坏变好,或者好变坏时,及时地调用各校级数据库服务器上的更新服务器状态服务;在状态不发生突变,无须及时通知的情况下,在同步服务器的作业队列队尾加上当前时刻服务器状态同步作业,目的是使各校级节点获取当前时刻资源服务器和资源镜像服务器运行状态。

3.4.4 作业队列工作算法

(1)查找是否有正在处理的任务,如果有,这个周期不处理,转(3);如果没有,转(2)。

(2)取队首作业,标记其状态为正在处理,调用相对应的处理逻辑。例如:对于工作内容为获取服务器状态,则根据所要获取的服务器的 IP 或者域名,调用该服务器发布的获取网格节点运行状态服务,获取成功标记该项作业成功,出现异常,标记异常同时记录异常信息。对于成功完成的作业,从该队列删除;对于出现异常的作业,从队首删除,并插入异常队列列表。

(3)退出。

3.4.5 作业异常的处理

对于作业异常的情况,管理员需根据具体的异常信息,进行故障排查后,对该项作业进行重做。

4 结束语

文中提出的方案,对已有系统修改较少,基本不增加原有系统负担,动态扩展性好,不影响原有系统的正常运行,能线性地提高并发度和系统的吞吐量,实践证明,该方案是有效的。对该方案还可以从以下几方面进一步完善:

(1)监控的周期参数,本原型系统固定的设置了监控周期,在网格环境变化的情况下,如何针对特定的环境,设置该参数,在不增加系统负担的情况下,更加有效地发挥系统的效率。

(2)分发策略,本原型系统在监控的状态发生通断状态时,才采取及时通知,下一步有必要对系统的压力进行测试,定义合适的阈值,当某个或某些参数达到阈值时采取及时通知。

(3)访问策略,校级服务器在选择最佳的服务器进行访问时,可以有多种方式,比如,同网络类型(教育

(下转第 55 页)

5 结束语

文中实验表明, Bagging 算法不能保证使单 SVR 的泛化能力得到提升。根据 Valentini 提出的 SVM 异构集成思想, 将 GA 引入用以搜索 SVR 的“低偏差区域”, 给出了一种基于 GA 的 SVR 异构集成方法, 有效提高了单 SVR 的泛化能力。在水质遥感监测的应用中, 当面临小样本情况时, 对水质参数的回归预测, 主要关心的是模型的预测精度。这种基于 GA 的 SVR 异构集成方法虽然比单 SVR 要费时些, 但在合理范围内付出较大的时间花销, 得到更精确的回归预测结果正是它的意义所在。

参考文献:

- [1] 杨一鹏, 王 桥, 王文杰, 等. 水质遥感监测技术研究进展[J]. 地理与地理信息科学, 2004, 20(6): 7-12.
- [2] 刘灿德, 何报寅. 水质遥感监测研究进展[J]. 世界科技研究与发展, 2005, 27(5): 40-44.
- [3] Vapnik V N. Estimation of Dependencies Based on Empirical Data[M]. Berlin: Springer-Verlag, 1982.
- [4] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
- [5] Kim H, Pang S, Je H, et al. Constructing Support Vector Machine Ensemble[J]. Pattern Recognition, 2003, 36(12): 2757-2767.
- [6] Valentini G, Muselli M, Ruffino F. Bagged Ensembles of SVMs for Gene Expression Data Analysis[C]//IEEE International Joint Conference on Neural Networks. Portland: [s. n.], 2003: 1844-1849.
- [7] Valentini G, Dietterich T. Bias-Variance Analysis of Support Vector Machine for the Development of SVM-Based Ensemble Methods[J]. Journal of Machine Learning Research, 2004(5): 725-775.
- [8] Smola A J, Scholkopf B. A tutorial on support vector regression[J]. Statistics and Computing, 2004, 14(3): 199-222.
- [9] 周兆永, 汪西莉, 曹艳龙. 基于 GA 优选参数的 SVM 水质评价方法研究[J]. 计算机工程与应用, 2008, 44(4): 190-193.
- [10] Dietterich T G. Machine learning research: four current directions[J]. AI Magazine, 1997, 18(4): 97-136.
- [11] Breiman L. Bagging Predictors[J]. Machine Learning, 1996(2): 123-140.
- [12] Dong Yan-shi, Han Ke-song. A Comparison of Several Ensemble Methods for Text Categorization[C]//IEEE International Conference on Services Computing: [s. l.]: [s. n.], 2004: 419-422.
- [13] Domingos P. A Unified Bias-Variance Decomposition[R]. Seattle: Department of Computer Science and Engineering, University of Washington, 2000.
- [14] Breiman L. Random Forests[J]. Machine Learning, 2001(1): 5-32.
- [15] Gunn S. Support Vector Machines for Classification and Regression[R]. Southampton: University of Southampton, 1998.
- [16] Houck C, Joines J, Kay M. A genetic algorithm for function optimization: a Matlab implementation[R]. North Carolina: North Carolina State University, 1995.
- [17] Zhou Zhi-hua, Wu Jian-xin, Tang Wei. Ensembling neural networks: many could be better than all[J]. Artificial Intelligence, 2002, 137(1-2): 239-263.

(上接第 51 页)

网、电信网等)的优先访问、负载最轻的优先、网络路径最短等等。建立最佳服务器的评价模型。

参考文献:

- [1] 张裔智, 王 磊. 基于 CDN 内容管理网络优化研究[J]. 微计算机信息, 2008, 24(6-1): 308-309.
- [2] 杨传栋, 余镇危, 王行刚, 等. 一种流媒体 CDN 的内容部分推送策略[J]. 计算机工程与应用, 2007, 43(25): 162-164.
- [3] 许东民, 任 宇. 基于智能 DNS 的策略路由在校园网中的实现[J]. 成都信息工程学院学报, 2007, 22(6): 716-718.
- [4] 宋兴彬, 许红果. 集群和负载均衡技术在省级数据集中的应用研究[J]. 计算机工程与设计, 2008, 29(1): 92-94.
- [5] Foster I. A Globus Primer or, Everything Your Wanted to Know About Globus, but were Afraid to Ask Describing Globus Toolkit Version 4[EB/OL]. 2005-08-05. <http://globus.org/toolkit/docs/4.0/key/>, 5/8/2005.
- [6] Foster I, Carl K. The Grid: Blueprint for a New Computing Infrastructure[M]. Second Edition. USA: Morgan Kaufmann, 2005.
- [7] Czajkowski K, Fitzgerald S, Foster I, et al. Grid information service for distributed resource sharing[C]//Proceedings of the 10th IEEE HPDC. San Francisco, CA, USA: [s. n.], 2001: 181-194.
- [8] Yeong W, Howes T, Kille S. Lightweight directory access protocol. RFC 1777[S/OL]. 1995. <http://www.ietf.org/rfc/rfc1777.txt>.
- [9] 杜艳明. 面向服务的体系架构与 Web 服务应用研究[D]. 武汉: 武汉科技大学, 2004.
- [10] 许 骏, 史美林, 李玉顺, 等. 网格计算与 e-Learning Grid[M]. 北京: 科学出版社, 2005: 16-20.
- [11] 王 琴, 曾文华. 网格资源调度算法的负载均衡及性能分析[J]. 微电子学与计算机, 2006, 23(10): 201-203.