

多用户网格资源管理系统及其在智能建筑中的应用

王 征¹, 刘 扬²

(1. 西安石油大学计算机学院, 陕西 西安 710065; 2. 河南工业大学信息科学与工程学院, 河南 郑州 450001)

摘 要: 网格资源管理系统的核心任务就是要聚合跨域异构、动态、增长的资源, 为上层用户及应用提供满足需求的资源服务, 而现有的网格资源管理系统多是独占的: 一方面是将资源管理与负载管理紧密耦合在一起, 系统通常只能为单一应用负载服务; 另一方面是用户使用资源的不受约束和管制, 易于造成资源的匮乏或浪费. 本文提出了一种多用户协调共享的网格资源管理系统 MCSGRMS, 是一种面向效用的资源管理系统, 支持多种用户和应用负载调度器对资源的协调共享, 改善了独占系统的可扩展服务能力. 将该系统应用于“智能建筑”和“智能社区”可以极大地降低整个社区企业的 IT 总成本.

关键词: 网格; 资源管理; 协调共享; 智能建筑

中图分类号: TP 393

文献标志码: A

文章编号: 1006-7930(2013)06-0891-05

网格是构建在 Internet 之上的一个开放的分布式系统, 其目标就是要在分布、异构、自治的资源环境中构造动态的虚拟组织, 实现跨越多个自治域的协同资源共享与资源协作, 在有限的时间内为网格用户或应用提供随需的聚合资源, 解决多样性的用户应用问题^[1]. 网格资源管理系统所关心的核心问题已经由“如何在所有的资源上快速运行一个(或一批)作业?”转变为“如何能为更多的用户提供按需服务, 以提高用户的满意度并获取更多的资源收益?”. 因而, 网格资源管理系统要首先能够为聚合 Internet 范围的跨越多管理域的各种资源, 并且协调利用所聚合的资源, 为大量用户应用提供一致的、动态部署的、高效协同的资源服务.

回顾现有的一些网格资源管理系统, 如 Globus^[2]、Legion^[3]、Nimrod-G^[4]、AppLes^[5]、Netsolve^[6]等, 这些系统通常都是以资源为中心, 它们存在的一些问题主要体现在以下几个方面: (1) 将负载管理和资源管理紧密耦合在一起, 导致网格中的资源只能为某种单一应用负载服务. (2) 在资源分配时很少考虑将资源分配给最需要该资源、最能产生效益的用户及其应用, 这样常常会导致资源分配的不合理和不公平, 导致用户应用的性能下降和用户的满意度下降. (3) 认为资源的提供是自愿和无条件的, 资源的使用是按需并且免费的, 这种资源共享方式在早期的分布式系统中是合理的, 因拥有者和使用者通常属于同一个组织, 但它不适用于网格, 专属的(dedicated)资源利用方式易导致资源使用方的需求膨胀和资源提供方的热情不足, 导致实用价值和应用前景降低.

提出并实现了一种多用户协调共享的网格资源管理系统 MCSGRMS, 致力于解决多种用户应用负载对资源的协调共享问题, 实现面向效用的资源管理系统^[7]. 第1节详细阐述了所提出的网格资源管理系统 MCSGRMS, 包括设计思想、体系结构和系统实现. 第2节给出了 MCSGRMS 的一个 SOA 应用实例, 给出了应用在该原型系统上的测试效果, 证实了该系统的有效性. 最后对全文进行了总结.

1 多用户协调共享的网格资源管理系统 MCSGRMS

1.1 设计思想

MCSGRMS(Multi-users coordinated sharing grid resource management system)是一种面向消费者共享的网格资源管理平台, 是一个较为通用的计算资源管理基础设施, 多种多样的用户应用可以同时运行在该平台上, 共享使用网格中的资源, 其上可以运行各种负载调度器或负载管理系统, 支持包括高

性能应用、批处理应用、SOA 应用、J2EE/.NET 等多种应用. 其核心思想及特征有以下几点:

(1) 解耦资源管理与负载管理, 使得资源管理层的功能接口更加清晰和明确, 不再像传统的资源管理系统那样将资源管理和负载管理紧密耦合在一起.

(2) 多种多样的用户应用, 抽象化为资源消费者, 支持基于资源消费者到资源提供者的协同共享.

(3) 用户的资源共享和分配是基于策略的, 中心协调者会按照一定的策略约束, 为消费者分配满足需求的资源.

(4) 支持跨越多组织的资源发现和监控, 可以通过使用 MCSGRMS 提供的命令行接口或 API 访问接口直接发现和获取资源.

(5) 简化了生命周期管理, 各种负载管理系统可以在 MCSGRMS 之上并存, 共享使用由 MCSGRMS 提供的资源.

(6) 提供一组可扩展的服务, 多样性的用户应用在这组服务之上享用资源, 而异构自治的资源在这组服务之下被虚拟化和聚合.

如图 1 所示, 解耦之后的 MCSGRMS 专门面向资源管理, 实施资源供应方的策略, 主要关心资源的优化使用等问题, 如哪些资源在什么时候可用, 哪个用户能够获得多少资源, 当资源不可用时如何配置额外资源等. 而运行在 MCSGRMS 之上的各种应用负载管理者则直接从 MCSGRMS 中获取资源, 主要关心应用负载的调度问题, 如哪个工作单元先执行, 哪个工作单元运行在哪个符合条件的资源上, 如何保证用户应用的 QoS 目标(响应时间、吞吐量、延迟时间)等.

1.2 体系结构

网格资源管理系统的体系结构如图 2 所示, MCSGRMS 位于本地资源管理者层与应用负载管理者层之间, 跨越了体系结构的三个层次: 管理服务层、扩展服务层和 API/SDK 层. 管理服务层包括提供者 Agent 和协调者 Agent; 扩展服务层包括消费者 Agent、事件服务、门户服务、安全服务、记帐服务、容错服务等一组软件部件(图中只显示出一部分软件部件); API/SDK 层提供了相应服务的功能接口 API. 位于 MCSGRMS 之下的包括两层, 分别是网格设备资源层和本地资源管理者层. 位于 MCSGRMS 之上的也包括两层, 分别是应用管理者层和用户应用层. MCSGRMS 能够支持的应用有很多, 各种应用管理者运行在 MCSGRMS 之上, 通过 API 或者服务接口获取网格中的资源, 通过自身调度供上层用户使用.

MCSGRMS 的核心功能有三个, 分别是发现(Discover)、分配(Allocate)和执行(Execute), 这三大核心功能分别对应网格资源管理的三个阶段, 第一阶段发现阶段, 为用户找到可用资源; 第二阶段分配阶段, 按照共享策略为用户分配资源; 第三阶段执行阶段, 在所分配的资源上执行用户应用, 记录资源使用情况, 释放资源.

MCSGRMS 采取的是一种可扩展的资源管理体系结构, 对资源的虚拟化采用 plug-in 的方式, 即每种类型资源的虚拟化功能对应一个 plug-in 插件, 系统通过加载各种 plug-in 来管理各种类型的资源, 核心功能部件能够兼容完成各种类型资源的发现、分配和管理任务. 目前, 在 MCSGRMS 中只实现了计算资源的 plug-in, 该 plug-in 负责完成计算资源的虚拟化任务, 通过将计算资源抽象成可共享使用的计算槽(slots), 支持面向 SOA/J2EE 此类事务型的应用, 即用户应用可建模成一个服务(service), 一个服务

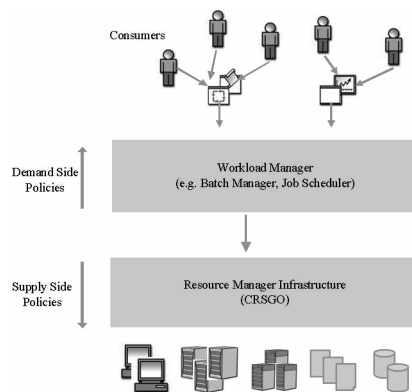


图 1 解耦资源管理与负载管理的设计思想

Fig. 1 The desingning idea of decoupling resource management and workload management

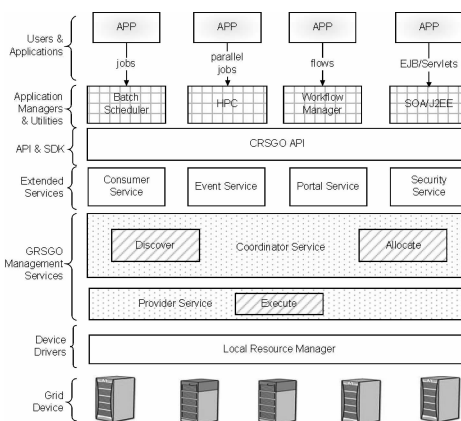


图 2 MCSGRMS 的体系结构

Fig. 2 The system structure of MCSGRMS

可以有若干服务实例(service instance)运行在网格计算资源上.

1.3 系统实现

MCSGRMS 的核心软件部件包括三部分,协调者代理服务(Coordinator agent)、提供者代理服务(Provider agent)和消费者代理服务(Consumer agent),其中消费者代理服务虽然处于扩展服务层,但是由于它能够大大方便上层用户使用网格,因而也将其并入核心部件. MCSGRMS 的核心框架结构如图3所示.

消费者代理(Consumer agent, CNA)负责为用户应用请求资源并且执行用户应用,主要包括应用需求规划器(Requirement Planner)和服务控制器(Service controller)两个子部件,应用需求规划器负责根据用户应用的特征属性和执行要求生成相应的资源需求,并将资源请求发送给协调者代理.支持 SOA/J2EE 等应用负载管理器,如 Websphere^[8]、Symphony^[9],即每一个用户应用对应一个服务,服务有一个相关的配置文件,采用 XML 表示方式,全面配置服务的属性特征、控制策略和分配约束等.应用需求规划器会加载服务的配置文件,并根据服务的模型参数生成相应的资源请求提交给协调者代理.服务控制器负责控制服务的执行,可以对服务及服务实例进行注册、启动、终止、更新等一些基本的控制操作.

协调者代理(Coordinator agent, CDA)负责为用户应用分配资源,在分配之前首先要发现网格中的可用资源,包括两个子部件,分别是资源分配引擎(Allocation engine)和资源信息发现者(Information discoverer).协调者代理运行在网格的虚拟组织中,可以聚合网格中的资源提供者,并根据共享策略约束统一为用户分配资源.资源分配引擎接收消费者代理发来的资源分配请求,周期性地分配资源,并且将分配结果通知给消费者代理,伴随着分配过程,也会有一些消费者的资源会被收回,收回资源的通知也会相应地发给消费者代理,提交分配请求与收到分配(收回结果)通知是一个异步过程.资源信息发现者负责聚合并发现网格中的可用资源,接收消费者代理或内部分配引擎发来的资源发现请求,并与内部资源分配引擎交互,为分配引擎选择合适的资源,帮助分配引擎做出分配决策.

提供者代理(Provider agent, PRA)负责管理所拥有的资源并且执行用户应用,包括两个功能子部件,分别是资源代理(Resource Broker)和用户应用执行机(Execution machine).虚拟组织中有多个提供者代理,每个提供者代理对外提供虚拟化的共享资源能力,满足一定的共享策略,供消费者使用.资源代理负责监控资源状态,并发布资源信息到协调者代理中.资源代理的任务主要有三项:虚拟化资源、监控资源状态、发布资源信息到协调者代理.目前 MCSGRMS 只实现了一种计算资源类型的 plug-in,该 plug-in 提供了对计算资源的访问监控接口,对资源的虚拟化和监控任务都是通过先加载 plug-in,然后利用 plug-in 所提供的功能接口来完成.资源代理负责将所拥有的可用资源信息发布到协调者代理,同时也接收来自协调者代理发来的查询请求,将当前的资源情况汇报给协调者代理.

用户应用执行机接收消费者代理发来的执行请求,经过认证安全检查之后,为服务实例构建执行环境,将其封装成一个活动,运行在分配的计算资源槽上,控制着活动的整个生命周期过程.用户应用执行机的主要功能包括:控制活动、监控活动并且向消费者代理汇报活动状态、收集活动运行时的资源使用状况等.

2 MCSGRMS 的 SOA 应用实例

MCSGRMS 采用解耦资源管理与负载管理的设计思想,实现了一种较为通用的网格资源管理平台,多负载用户能够协同共享网格中的资源,完成用户应用的需要.本节介绍的 SOA(Service-oriented architecture)即面向服务的体系结构,应用负载调度器(Scheduler symphony)构建在 MCSGRMS 之上的应用案例,并给出了不同配置情况下的性能测试结果. SOA 是一种面向服务的应用形式,通过将企业

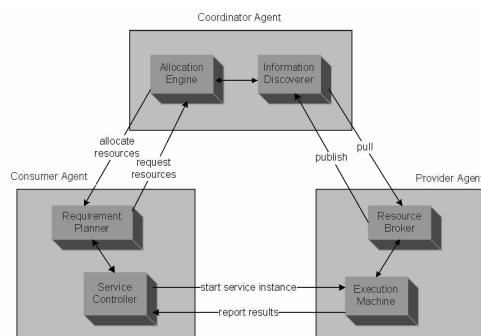


图3 MCSGRMS 核心框架结构图

Fig. 3 The structure of the core framework of MCSGRMS

逻辑封装成一个或多个服务来完成用户应用. SOA 应用负载调度器,如 Symphony 负责完成请求(message)到服务的映射和调度工作.

相对传统的构建于集群之上的应用负载调度器而言,构建在 MCSGRMS 之上的 Symphony 显然能够获得以下三个典型好处:(1)能够根据应用需求相应地获取网格中的资源,从而按需地启动服务实例,实现服务对用户请求的快速响应.(2)通过 MCSGRMS 提供的服务实例重启动(restart)功能,能够实现服务的高可用性.(3)利用 MCSGRMS 的资源重分配(reallocate)功能,能够实现服务的可扩展性.

Symphony 从 MCSGRMS 中获得分配资源之后启动服务,服务启动后形成会话(session),一个会话中包含一组任务(Task),任务的输入即为一个请求消息(message),任务执行完成后产生输出.图4用图表方式显示出一个 SOA 用户应用的执行状况,包括已经关闭的会话(Closed sessions)情况,异常中止的会话情况(Aborted sessions),已经完成的任務(Done Tasks)情况和发生错误的任务(Error Tasks)情况.

我们对运行于 MCSGRMS 之上的 SOA 应用进行了系统集成测试,通过远程安装的方式将 MCSGRMS 一次性安装在 200 台装有 Pentium 2.6 GHz 处理器的工作站上,工作站采用 Windows 或者 Linux 操作系统.鉴于备份/恢复(Recoverable)SOA 任务对应用的性能影响较大,因而在测试时考虑了该参数.我们从任务的执行时间(Task run time)和任务的投交吞吐量(Task sending throughput)两方面来考察 SOA 应用的执行情况,从 CPU 效率(CPU Efficiency)来考察 MCSGRMS 的资源利用率情况.系统集成测试的性能结果见表1.

通过测试可以看出,MCSGRMS 的资源利用率一直保持在较高的水平,在 200 台 CPU 的测试条件下,平均资源利用率达到 99.64%,并且 SOA 应用的任务吞吐量和任务执行时间都得到了明显改善.

目前该系统可以应用在将云计算和物联网技术相融合的“智能建筑”之中,进而构成“云物”缭绕的智能社区.在“云物”缭绕的智能社区里,每一栋智能化的建筑,无论是住宅、学校、办公楼、医院等等,都处在基于互联网的“云物”包围中,使得社区里的人们对信息的掌握变为了对于信息的有效利用,使得信息能够对社区的电力设备、医疗机构、办公机构、学校、生活娱乐设施等产生有利地直接影响.据估算,2012 年我国的“智能建筑”集成系统市场规模已经从 2005 年的 201 亿元人民币跃升至 2012 年的 1 223 亿元人民币,预计到 2015 年将接近 3 000 亿元人民币.随着云计算和物联网技术与智能化建筑的不断融合,“云物”型智慧社区的比例势必会逐年增加.可以预见,我国未来在智能化建筑和智能化社区方面将会有长足的发展,采用本系统将帮助社区中企业的 IT 人员即时响应不断变化的业务需求,对资源进行实时分配与重分配;极大的提高 IT 资源的分配效率以及资源的可扩展服务能力,从而降低了整个智能化社区企业的 IT 总成本.

3 总 结

本文提出并实现了一种多用户协调共享的网格资源管理系统 MCSGRMS,可提供虚拟化及自动化功能,为各种应用类型提供一个共享所有 IT 资源的网格平台,是一种较为通用的网格资源管理基础设施.在实践应用中显示,该系统较之以往的网格资源管理系统,能够支持的用户应用种类更多,高性能应用、批处理应用、SOA 应用等多种应用模式可同时运行在 MCSGRMS 之上,并且资源的分配效率更高,

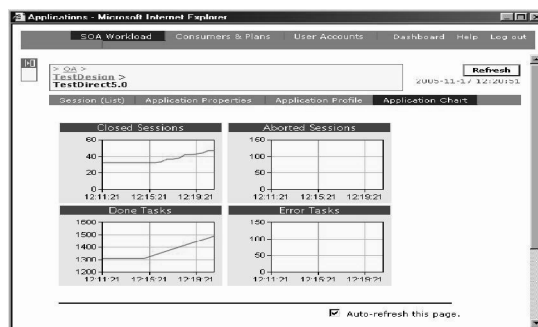


图4 构建在 MCSGRMS 之上的
SOA 应用负载运行状况

Fig. 4 The load operation situations of SOA application based on MCSGRMS

表1 200 CPU 性能测试结果

Tab. 1 The performance test result of 200 CPU

操作系统	备份/恢复	任务执行时间	CPU 效率	任务投交吞吐量
Windows	Yes	1 sec	99.56%	1 047
Windows	No	1 sec	99.54%	1 775
Linux	Yes	1 sec	99.75%	2 000
Linux	No	1 sec	99.72%	2 914

因而用户的满意度也得到了明显提高. 实践应用显示, 该系统较之以前的网格资源管理系统, 具有支持的应用种类多, 资源的分配效率高等优点, 资源的可扩展服务能力得到了有效改善, 因而用户的满意度也得到了明显提高. 将该系统应用于“智能建筑”和“智能社区”中, 可以极大地降低整个智能化社区企业的 IT 总成本.

参考文献 References

- [1] CHAPIN S, CLEMENT M, SNELL Q. A Grid Resource Management Architecture, Strawman 1[M]. Manchester: Grid Forum Scheduling Working Group, 1999.
- [2] FOSTER D, KESSELMAN C. Globus: A metacomputing infrastructure toolkit[J]. International Journal of Supercomputer Applications, 1997, 11(2): 115-128.
- [3] GRIMSHAW F, WULF A. The Legion vision of a worldwide virtual computer[J]. Communications of the ACM, 1997, 40(1): 39-45.
- [4] BUYYA R, ABRAMSON D, GIDDY J. Nimrod-G: Architecture for a Resource Management and Scheduling Systems in a Global Computational Grid[C]//The 4th International Conference on High Performance Computing in Asia-Pacific Region, Beijing, 2000: 283-289.
- [5] BERMAN F, WOLSKI R, FIGUEIRA S, et al. Application level scheduling on distributed heterogeneous networks[J]. Proceedings of Supercomputing, 1996, 2(3): 111-116.
- [6] CASANOVA H, DONGARRA J. Net Solve: A networks server for solving computational science problems[J]. International Journal of Supercomputing Applications and High Performance Computing, 1997, 11(3): 212 - 223.
- [7] PADALAP, KANG G. Adaptive Control of Virtualized Resources in Utility Computing Environments [J]. EuroSys'07, 2007, 6(5): 89-95.
- [8] Websphere[EB/OL] <http://www.ibm.com/developerworks/cn/websphere/geoconst.html>.
- [9] Platform Symphony: Enterprise Grid Solutions for Financial Services[EB/OL] http://www.platform.com/NR/rdonlyres/4EFB903A-3AED-485A-90DA-B87E5D7DAB57/0/Plt_Symphony3_Br.pdf/202258.html.

Multi-users grid resource management system and its applications in intelligent buildings

WANG Zheng¹, LIU Yang²

(1. College of Computer Science, Xi'an Shiyou University, Xi'an 710065, China;

2. College of Information Science and engineering, Henan University of Technology, Zhengzhou 450001, China)

Abstract: The key task of grid resource management system is to supply on-demand resources service for various users and applications by aggregate heterogeneous, dynamic and increase resources. Most current grid resource management system is exclusive bringing two sides of deficiency distinct. On the one side, resource management and workload management are tightly coupled, resulting in the singleness of application supported. On the other, applications can get resources with no economic cost or payment and utilization is less considered and controlled to support multi applications, which results in the scarcity or waste of resources. We bring forward and realize a multi-user coordinated sharing resource management system, which is a utility oriented resource management system to support resources sharing by multi users or multi workload schedulers, so extendable resource serving capability can be improved. When the system is used in “intelligent building” and “intelligent community”, the IT cost of the community enterprises can be reduced significantly.

Key words: *grid; resource management; coordinated sharing; intelligent building*