

# 基于GMA的层次化网格资源信息服务研究与实现

张瑞雪, 陈 萃, 李 军, 祁 羽

(国防科学技术大学电子科学与工程学院, 长沙 410073)

**摘要:** 网格资源信息服务(GRIS)是网格基础结构的关键组成部分, 它定义了一套发现和监控网格资源的机制, 并以此为网格上层应用提供服务, 如故障检测、性能分析、任务调度等。资源信息的准确获取、有序管理和高效检索是其中的关键技术问题。基于GMA体系结构, 文章提出了一种对网格资源信息进行层次化组织管理的方法, 并面向Windows平台, 设计实现了基于该方法的网格资源信息服务系统。论文研究成果已在工程项目中得到实际应用。

**关键词:** 网格; 资源信息服务; GMA

## Research and Implementation of A GMA-based Hierarchical Grid Resource Information Service

ZHANG Ruixue, CHEN Luo, LI Jun, QI Yu

(Dept. of Electronic Science & Engineering, National University of Defence Technology, Changsha 410073)

**【Abstract】** Grid resource information service (GRIS) is one of the key parts of grid infrastructure. It defines a set of mechanisms for discovery and monitoring of grid resources, and thus provides information for higher-level grid application, such as fault detection, performance analysis, and scheduling etc. Accurate obtaining, effective managing and efficient retrieving of resource information are the key technologies in GRIS. This paper proposes an approach which is based on GMA architecture and hierarchically organizes grid resource information. A Windows-oriented GRIS is designed and implemented with this approach, and it has been applied in a project.

**【Key words】** Grid; Resource information service; GMA

“网格”就是在动态形成的虚拟组织之间灵活、安全、协同地共享资源, 而网格资源信息服务是实现网格应用的关键<sup>[1]</sup>。资源信息服务的主要作用是发现资源并对资源进行监控, 进而为上层应用提供资源信息以完成各种任务, 如故障检测、性能分析、性能调节、性能预测、任务调度<sup>[2]</sup>。如何在分布、异构的环境下实现对资源信息的准确获取、有序管理和高效检索, 是资源信息服务需要解决的关键技术问题。

目前传统的操作系统(如Windows NT/2k和Unix)中都提供了完备的单机状态和性能监控组件, 但都不支持远程访问。SNMP协议能够提供远程访问的功能, 但只提供了点对点的访问机制。集群系统中的监控组件虽然是分布式的, 但监控对象的地理位置相对集中, 结构也基本相同<sup>[4]</sup>。因此上述传统的监控方法不能适应网格资源信息服务的要求。

本文分析了网格资源信息服务的基本需求, 并面向Windows平台, 基于GMA体系结构提出一种层次化管理分布式资源信息的方法。该方法能够方便地实现资源信息的发现、获取、归档、检索等核心功能, 同时具有很好的可扩展性。

### 1 相关研究

#### 1.1 研究现状

目前, 国内外一些研究机构已经开发出了若干网格资源信息服务系统, 其中比较著名的有: R-GMA(Relational Grid Monitoring Architecture), MDS(Monitoring and Discovery Service)。国内的有: GridMon<sup>[4]</sup>, GridEye<sup>[5]</sup>等。R-GMA的特点是使用了关系模型, 使得每个虚拟组织看上去都有一个大

的关系数据库一样。MDS是Globus Toolkit的一个组件, 它构建在LDAP协议的基础之上, 为访问由底层信息提供者收集的性能数据提供了统一、灵活的接口。另外, 由于它采用了分散式的结构, 因此具备了较强的可扩展性。国内的GridMon和GridEye借鉴了众多资源信息服务系统的优点, 较好地满足了资源信息服务的需求。

本文研究的目的是实现一个基于GMA体系结构的层次化资源信息服务系统模型。

#### 1.2 GMA 体系结构

网格资源信息服务系统必须满足低延迟、高数据流量、对被监控资源的影响最小、安全和具有可扩展性等需求<sup>[2]</sup>。目前GGF提出的GMA体系结构已经被确认为网格资源信息服务系统设计与实现的标准。

在GMA中, 资源信息被作为一种带有时间戳的事件(event)来处理, 事件在生产者(producer)、消费者(consumer)和目录服务(directory service)这3种类型的组件之间流动, 构成了信息的交换。生产者事件源, 它响应并传送事件数据给消费者。消费者用来请求和接收事件数据。生产者和消费者都把自己的信息注册到目录服务当中, 以便能够互相发现。目录服务只保存性能事件的元数据以及这些事件到相关的生

**基金项目:** 国家“863”计划基金资助项目(2003AA135110)

**作者简介:** 张瑞雪(1977—), 男, 硕士生, 主研方向: 信息集成; 陈萃、李军, 博士、副教授; 祁羽, 硕士

**收稿日期:** 2005-11-11 **E-mail:** 13487318414@humcc.com

生产者或消费者的映射，而并不保存事件数据本身。

GMA 体系结构实现了数据发现与传输的分离，但是它并没有提出具体的解决方案。为此我们基于 GMA 体系结构，设计实现了一个层次化的网格资源信息服务系统模型。

## 2 层次化网格资源信息服务

### 2.1 资源信息的模型和表示

网格资源是一个抽象的概念，它可以是硬件(如主机、设备)，也可以是软件(如操作系统、应用程序)。在计算机系统中，处理器、存储器、网络及其所运行的应用程序是影响系统性能的主要因素，系统性能调整也主要在于如何在这些因素中获得某种平衡。因此将这些因素作为主要监控目标，把资源的信息模型定义为静态和动态两类。其中静态信息是指那些生命周期很长的资源信息，如计算机的 CPU 主频、内存大小、操作系统类型。而动态信息是指那些生命周期短，更新频率快的资源信息，如计算机的 CPU 使用率、内存使用率、当前网络负载、进程或服务的运行状态。

为了方便用户对节点性能信息的查询，我们还基于 Web Service 技术开发了统一的访问接口。用户通过调用 Web 服务，可以得到一个包含了节点性能信息的 XML 文档。

### 2.2 体系结构

资源信息服务的关键技术问题是能准确及时地获取资源信息并对之进行有效的管理与检索，这也是在设计系统时重点考虑的。系统的具体设计如图 1 所示。

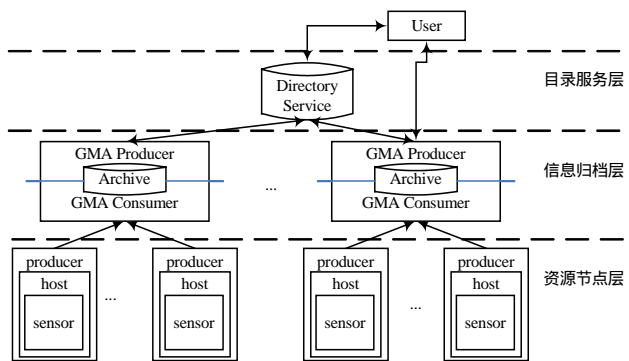


图 1 系统体系结构

整个系统分为 3 层：分别是资源节点层、信息归档层和目录服务层。其中，目录服务层也是分层次组织和管理的，这点将在 2.4 节中详细介绍。下面参照图 1 对系统的工作流程作一个简单的说明：在资源节点层中，每一个被监控节点(Host1、Host2、……)上都部署一个传感器(sensor)，用以收集静态和动态资源信息。这保证了资源信息获取的准确与实时性。传感器定时将采集到的信息交给生产者，由生产者组织成可理解的格式，称为性能事件。随后生产者将性能事件传递给信息归档层的归档系统(Archive)。归档系统在这里既是一个消费者也是一个生产者。作为消费者，它从生产者那里接收事件数据，并将之存储到关系型数据库中；而作为生产者它又可以响应消费者对归档事件的查询，返回结果事件集。归档系统能够将元数据信息发布到目录服务层，如它所存储的事件的类型以及事件与生产者的映射等。目录服务也采用关系型数据库来存储数据。使用者(user)如希望查询某个节点的性能事件，先要到目录服务中去查询这个节点属于哪个 Archive，然后再向相应的 Archive 发出请求。采用成熟的关系型数据库技术来组织和管理资源信息数据能够提高系统的有效性和可靠性。同时，采用分层组织和管理的办法也

符合实际当中一般单位或机构的管理模式，因此加强了系统的可扩展性和易管理性。

### 2.3 系统主要组件

#### 2.3.1 Sensor

传感器(Sensor)用于收集资源信息，它的设计要遵循以下原则<sup>[3]</sup>：轻量级(也就是传感器要设计的小巧、简洁)，对被监控资源的干扰小、低延迟、远程可访问、可移植性强等。

因为在实际应用中绝大多数节点计算机所运行的都是 Windows 系列的操作系统，我们着重面向 Windows 操作系统，开发了基于 WMI(Windows Management Instrumentation)技术的传感器。WMI 是微软提出的 Windows 环境自动管理方案。它采用统一的、基于标准的、可扩展的、面向对象接口，提供了与系统管理信息交互的标准方法。由于使用了 WMI 技术，因此我们的传感器具备了小型化、对被监控对象的资源占用率低，信息采集准确、延迟小以及较强的移植性等优点。

#### 2.3.2 Archive

归档系统用于存储传感器收集到的资源信息。因为目前多数资源信息服务系统采用层次数据模型(如 LDAP)来存储资源信息有一定的局限性，所以改用关系型数据库，它可以保存资源信息完整的历史轨迹。使用者可以有 3 种不同的查询方式：(1)对资源的当前性能信息的查询(最新查询)；(2)从当前时刻起对资源性能信息进行连续查询(连续查询)；(3)对资源在过去一段时间内的性能信息的查询(历史查询)。从 Archive 获取性能信息相对于从生产者直接获取，所得到的查询结果肯定是“陈旧”的。但是只要传感器对性能信息的采集频率足够快，性能信息归档的时间延迟足够短，这种“陈旧”是可以忍受的。

如图 1 所示，我们在关系型数据库的基础上实现了基于 GMA 体系结构的消费者接口和生产者接口。消费者接口的作用是接收事件生产者产生的事件数据，并存储在 Archive 中，生产者接口的作用有两个：一是响应使用者的查询将关系数据库中有事件数据返回给使用者；二是将 Archive 的一些元数据信息注册到目录服务器中去。

在本系统中，一般是在地理集中或属于同一组织机构的节点中配置一个 Archive。采用这样的配置方式可以使资源信息的采集和传输获得可靠和低延迟的保障。再者，如果处于同一个机构之内，在信息传输方面也可以免去企业级防火墙的干扰。

由于事件数据会不断地向 Archive 存储，因此 Archive 数据库的大小可能会无限制地增大。为了防止这种情况的发生，可以采取两种维护机制：一是限定数据库的大小，当数据库的容量超过某个阈值时，清除数据库中的陈旧数据；二是定时清除陈旧的数据。这两种机制都可以保证数据库的体积不会过于庞大。

#### 2.3.3 Directory Service

目录服务是资源信息服务的核心组件。本系统中仍使用关系型数据库来实现目录服务。使用标准的 SQL 语言来完成对目录服务的查询操作。目录服务的关系数据库中存放其下属 Archive 的注册信息，如每个 Archive 的位置，每个 Archive 包含哪些节点；每个节点的一些基本特征等。使用者如果想查询某个节点的某个性能指标，首先要查询目录服务以确定节点所属的 Archive 的地址，然后发出对目标节点性能数据的查询命令，Archive 接到查询后处理并将查询结果返回给使用者。

每个 Archive 在目录服务中对应的是一条记录。目录服务支持对记录的添加、更新、删除等操作。当有一个新的 Archive 向目录服务注册时,目录服务就相应地添加一条记录;当某个 Archive 的状态发生改变时,目录服务就更新相应的记录。目录服务中的每条记录都服从软状态生命周期管理。也就是每条记录都有一个生命期,如果在整个生命期之内都没有接到 Archive 的“keep-alive”信号,那么这条记录就被视为过期。目录服务会自动清除过期的记录。这种机制使得失效的 Archive 不会出现在目录服务的记录中。

#### 2.4 目录服务的分层管理与备份机制

网格中资源的数量众多,这是网格的一个特征。在这种情况下,单一的目录服务可能会因为所存储的数据量过于庞大而导致查询效率低下。另外从可靠性方面考虑,单一目录服务也会成为资源信息服务系统的性能瓶颈。因此有必要采取分层次的目录服务管理,这样一方面可以避免因单个目录服务器崩溃而导致整个网格资源信息服务系统失灵的情况,另一方面也更加符合实际应用中资源在一个组织机构中的组织方式。我们根据虚拟组织的概念来划分目录服务的管理范围。虚拟组织是为了完成某一个确切的业务而动态形成的资源的集合体。每个完成特定任务的虚拟组织中配置一个目录服务(可能还有一个备份目录服务)。多个虚拟组织又可以聚合成一个更大的虚拟组织以完成更复杂的任务,同时它们中的目录服务也聚合成更高层的目录服务。这个更大的虚拟组织可以称为“高层虚拟组织”,所拥有的目录服务称为“高层目录服务”,而组成它的小的虚拟组织可以称为“低层虚拟组织”,相应的目录服务称为“低层目录服务”。这就是目录服务的分层管理,可以用图 2 做一个简单的说明,图中,每个虚线方框部分划分为一个虚拟组织。

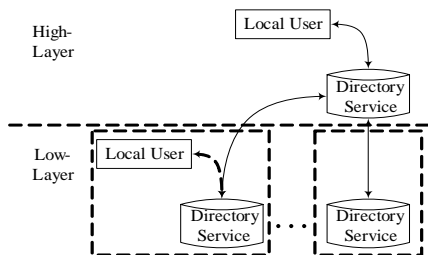


图 2 目录服务的分层管理

如图 2 所示,低层目录服务向高层目录服务注册,每个目录服务可以有自己本地用户,本地用户可以查询本地目录服务下所有的节点的性能信息。

目录服务是网格资源信息服务的核心组件,一方面使用者只有通过它才能发现所要的资源信息,另一方面资源也只有向它注册才能被使用者所发现。一旦目录服务器瘫痪或不能访问,那么使用者将无法发现和定位任何想要的性能数据。为了解决这个性能瓶颈问题,一方面可以采用分层次的目录服务管理,这样即使某个目录服务器出现了故障,其它目录

服务器也可以照常工作;另一方面,要建立备份机制。就是定期(或当目录服务的内容发生改变时)对目录服务的内容进行冗余备份。当主目录服务器发生故障时,备份目录服务器将接替主目录服务器的工作,所有用户的请求也将被重定向到备份目录服务器上。如果这时有新加入的 Archive(或低层目录服务),则将它注册到备份目录服务器上。待主目录服务器恢复正常后,主、备份目录服务器同步注册信息,然后重定向使用者的请求到主目录服务器上去。这样,只有当两台目录服务器同时出现故障时才会影响整个监控系统的正常运行。这样的概率是非常小的,因此采用这种备份方式能够大幅度提高整个网格监控系统的可靠性。

综上,本系统可以由下到上层次化为资源节点层、信息归档层、目录服务层,其中目录服务层又可以根据不同的虚拟组织规模细分为多级层次。至此,我们实现了一个基于 GMA 的层次化资源信息服务系统模型。网格资源信息服务系统采用这种层次化的管理模式,使得网格资源信息服务系统层次分明、便于管理,符合实际应用背景下计算节点的组织形式,且具有较强的可扩展性。

我们已应用上述的系统模型和工具,在面向国家地质调查(NGG)的应用中实现了一个原型应用系统——国家地质网格资源信息服务系统(NGGRIS)(<http://www.ngg.cgs.gov.cn/>)。其资源信息服务的核心即是层次化网格资源信息服务。

### 3 结论

在网格应用当中,为了聚合大量动态、自治和异构的资源,网格资源信息服务具有很关键的作用。本文参考了国内外一些网格资源信息服务的实现方案,结合 NGG 网格应用的实际情况,基于 GMA 体系结构开发了一个层次化网格资源信息服务的原型系统。实践证明,该系统运行情况稳定,信息采集准确,易于管理,查询方便,可扩展性好,能适应整个 NGG 网格应用的需求。今后将增加更多的实验节点,进一步完善目录服务的分层次管理和备份的机制。

#### 参考文献

- 1 Zhang Xuehai, Freschl J L, Schopf J M. A Performance Study of Monitoring and Information Services for Distributed Systems[C]. Proceedings of HPDC-12, 2003.
- 2 Tierney B, Aydt R, Gunter D, et al. A Grid Monitoring Architecture[Z]. <http://www.didc.lbl.gov/GGF-PERF/GMA-WG/papers/GWD-GP-16-1.pdf>.
- 3 Ribler R L, Vetter J S, Simitci H, et al. Autopilot: Adaptive Control of Distributed Applications[C]. Proceedings of the 7<sup>th</sup> IEEE Symposium on High Performance Distributed Computing, Chicago, 1998-07.
- 4 查礼,徐志伟,林国璋等.基于 LDAP 的网格监控系统[J].计算机研究与发展,2002,40(8):930-936.
- 5 褚瑞,肖依,卢锡城.一种基于 GMA 的开放式网格资源信息服务.计算机研究与发展,2004,42(12):2114-2122.

(上接第 102 页)

- 4 Greco S, Matarazzo B, Slowinski R. Rough Sets Methodology for Sorting Problems in Presence of Multiple Attributes and Criteria[J]. European Journal of Operational Research, 2002, 138(2): 247-259.

- 5 Greco S, Matarazzo B, Slowinski R. Rough Sets Theory for Multicriteria Decision Analysis[J]. European Journal of Operational Research, 2001, 129(1): 1-47.