

SIG 中的资源监控系统*

田鸣华¹, 肖 侗¹, 刘 波¹, 黄 斌²

(1. 国防科技大学 计算机学院, 湖南长沙 410073; 2. 怀化学院 计算机系, 湖南怀化 418008)

摘要: 空间信息网格(SIG)是一个广域分布的信息基础设施,发现与共享闲置资源是其主要功能之一。为了实现这一目标,必须实现资源监控系统来聚集和分发资源的状态信息。介绍了在SIG中实现的资源监控系统——SIG-RMS。该系统遵循GMA架构,实现了局部监控服务(LMS)和全局监控服务(GMS)。系统采用模块化设计,具有灵活的可配置性,可扩展性和高可用性。详细介绍了SIG-RMS模块化的体系结构、灵活可扩展的信息模板和推拉式的通信方式,并描述了它的实现,最后总结了SIG-RMS的特点。

关键词: 空间信息网格; 资源监控; GMA; 信息模型

中图法分类号: TP311.5 文献标识码: A 文章编号: 1001-3695(2005)07-0177-03

Resource Monitoring System in SIG

TIAN Ming-hua¹, XIAO Nong¹, LIU Bo¹, HUANG Bin²

(1. School of Computer Science, National University of Defence Technology, Changsha Hunan 410073, China; 2. Dept. of Computer, Huaihua College, Huaihua Hunan 418008, China)

Abstract: Spatial Information Grid is a wide area distributed information infrastructure. Discovering and sharing of spare resources is the key capability of this system, which relies on a resource monitoring system for gathering and distributing status information of resources. This paper presents a detailed implementation of a resource monitoring system in SIG: SIG-RMS. It follows the GMA framework and implements Local Monitoring Service(LMS) and Global Monitoring Service(GMS). This system adopts module-based designing strategies, which makes it flexibly configurable, scalable and high available. The paper introduces the architecture, Information Model and communication ways of the SIG-RMS, and describes its implementations, finally conclude its main characteristic.

Key words: Spatial Information Grid(SIG); Resource Monitoring; GMA; Information Model

1 引言

空间信息网格(Spatial Information Grid, SIG)是一种汇集和共享地理上广域分布的海量空间信息资源,对其进行一体化组织与协同处理,从而具有按需服务能力的空间信息基础设施。对于这样一个广域分布的信息基础设施来说,共享与充分挖掘闲置资源能力,提高资源利用率是其研究的主要目标之一。为了实现这一目标,通过资源监控聚集和分发资源的状态信息就变得非常重要。通过监控及时采集资源的状态信息,可以用来帮助分析系统性能,预测系统运行的趋势,为资源调度、作业性能优化等应用提供可靠依据。

目前,美国、欧洲和日本等国家纷纷开展网格资源监控技术的研究。为了促进网格资源监控系统开发的互操作性,全球最大的网格论坛组织GGF为此专门成立了网格性能监控小组,该小组专门针对网格监控研究提出一个指导性体系结构(Grid Monitoring Architecture, GMA)^[1]。当前基于GMA实现的网格资源监控系统主要有MDS^[2], R-GMA^[3]和NWS^[4]等。MDS3(Monitoring and Discovery Service)是Globus项目中嵌入

实现的资源监控与信息服务系统,系统与其他网格子系统(如GSI)紧密集成,正是对GT3容器的过分依赖限制了它的推广使用。NWS是一个网络资源监控工具,能够完成广域环境下部分资源状态信息(如CPU利用率、网络带宽、延迟等)收集,并且提供系统性能预测功能。由于监测信息不太完整、访问接口的相对复杂以及结构设计上的缺陷使它不能广泛应用。RGMA是相对来说功能比较完整,遵循GMA规范实现的网格资源监控系统,但是就目前来说,它对软件的过分依赖使得它的安装、使用和维护非常复杂,而且暂时它只局限于在欧洲数据网格(EDG)项目内部使用。

针对现有系统的不足和SIG系统的特点,在SIG项目中我们设计和实现了SIG-RMS:一个遵循GMA规范实现的资源监控系统。系统是一个建立在广域环境上的分布式系统,能够对SIG环境中所有节点进行状态监视和监测数据的管理。系统设计逻辑上分为局部监控服务(LMS)和全局监控服务(GMS)两个层次。系统通过功能模块化设计具有灵活的可配置性、可扩展性和高可用性。此外,通过对资源监测数据的分层管理最大限度减少用户访问监测信息所需的性能开销。

2 SIG-RMS 系统设计

SIG-RMS目标是设计一个广域环境下的分布式资源监控系统。系统要求能够及时采集、分析和处理广域网环境下异构的各种资源的状态信息,如计算、存储、网络等,并将它们统一

组织起来, 通过系统提供的访问接口和服务屏蔽资源的异构性, 为用户提供直观、一体化的局部或全局监控视图, 同时为其他网格服务(如资源调度)提供方便的基于 Web 的编程接口。为了保持与其他类似网格监控系统的可交互性, 我们遵循 GMA 规范设计, 并将系统设计为两个逻辑层次, 即局部监控域和全局管理域。局部监控域负责一个站点或区域的资源监控和局部资源信息的存储; 全局管理域负责对各局部监控域的监控和管理, 以及全局信息的存储。

2.1 系统体系结构

SIG-RMS 系统主要由四个部分组成: 传感器 (Sensor)、局部监控代理 (LMA)、局部监控服务 (LMS) 和全局监控服务 (GMS)。传感器主要负责对局部监控域内资源状态信息的获取, 并向局部监控代理提供所获取信息。局部监控代理部分的功能是将传感器提供的源信息进一步地挖掘处理并写入数据库作为历史记录。局部监控服务部分的功能是从局部监控数据库中取得资源信息, 接受局部范围内用户查询或者其他网格应用调用的合法请求, 并且通过 Web 模式发布监控信息。全局监控服务部分的功能是从全局监控数据库中取得资源信息, 接受广域范围内用户查询或者其他网格应用调用的合法请求, 并且通过 Web 模式发布监控信息。SIG-RMS 资源监控系统体系结构如图 1 所示。

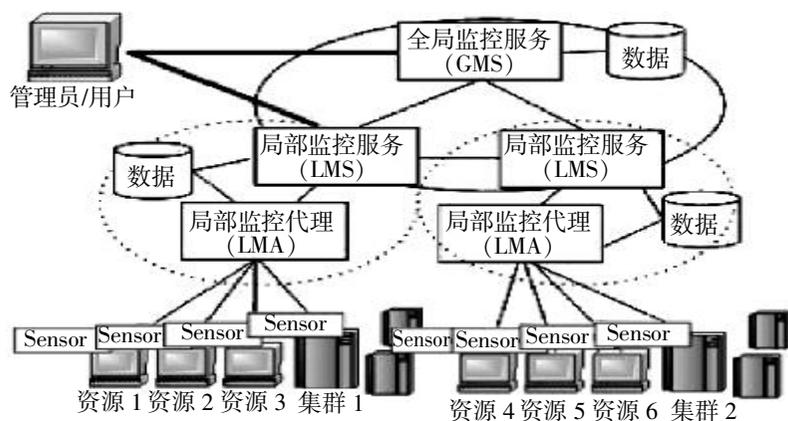


图 1 SIG-RMS 系统体系结构

(1) 传感器 (Sensor)。Sensor 是驻留在要监控的资源节点上(包括网络上的服务器、工作站以及集群计算节点)的进程。其功能是取得所在节点的监控源信息并对外提供信息。对网络上的服务器、工作站等资源节点的监视信息包括节点系统静态(如操作系统类型、CPU 主频、存储总容量、网络地址等)和动态(如处理器负载状况、可用存储空间、内存可用大小等)两类信息; 对集群资源节点的监视信息包括两个部分, 即集群整体计算能力信息和集群服务器当前运行状态信息。

(2) 局部监控代理 (LMA)。LMA 是运行在局部监控服务器上的一个守护进程。最初运行时需要向 LMS 进行注册。主要负责维护局部监控域内各个资源节点状态信息更新, 并且依次收集局部监控域内各个资源节点中 Sensor 提供的监测信息。LMA 还负责每次将收集的资源节点监测信息进行进一步的分析和处理, 然后将监测信息记入数据库, 作为所在监控域各资源节点性能的历史数据进行保存。这些历史数据是为以后进行各个局部监控域性能分析对比的重要依据。同时, LMA 还负责接受局部监控服务 (LMS) 访问实时资源监测信息的请求并作出相应。

(3) 局部监控服务 (LMS)。对应于 GMA 中的 Customer/Producer 复合组件, 最初运行时需要向 GMS 进行注册。它的主要功能是维护局部监控域内相应 LMA 的注册信息以及与它

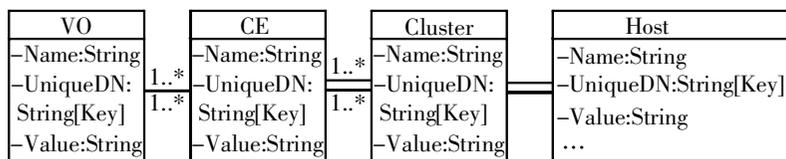
相邻域监控服务的相关信息; 从局部监控数据库中取得资源监测信息; 定期对所监控局部域内的资源监测信息生成详细监测报告, 为用户提供局部监控视图。同时接受用户查询或者其他网格应用调用的合法请求, 并且通过 Web 模式发布出监控信息。如果某个用户或网格应用要访问的监控数据不在本局部监控服务中, LMS 还负责查询 GMS 找到满足条件的 LMS, 然后转发请求, 获取相应监控数据, 最终返回给用户。

(4) 全局监控服务 (GMS)。维护多个局部监控域 LMS 相应的注册信息, 接受和响应 LMS 查询其他 LMS 相关信息的合法请求; 负责从各个局部监控数据库中抽取资源关键监测数据; 定期对所监控多个局部域内的资源监测信息生成总体监测报告, 为用户提供全局监控视图。同时接受用户查询或者其他网格应用调用的合法请求, 并且通过 Web 模式发布监控信息。

2.2 资源信息模型 (RIM)

网格环境下要监控资源的信息种类繁多, 其特征各不相同, 数量也不相同, 资源的状态动态变化。要把各类信息统一组织, 首先就需要建立统一的信息模型, 便于监测数据的统一管理和访问。资源信息模型是对于所有要存储信息的整体逻辑结构的设计, 应当具有良好的可扩展性、灵活性和较宽的适应面。资源信息模型应该关注要监控资源的状态以及它们之间关系的一致描述, 而不依赖某一具体技术的实现。

SIG-RMS 采用了一个简单的层次资源信息模型-RIM。该模型扩展了 Grid Laboratory Uniform Environment (GLUE) 发布的通用信息模板, 为 SIG-RMS 监控各种异构分布资源数据提供了一个一致的视图。RIM 为各种监控资源的状态信息以及它们的相互关系定义了一致的命名和结构属性。当前描述的监控资源包括: 虚拟组织 (VO)、计算节点 (CE)、主机 (Host)、集群 (Cluster)、处理器 (Processor) 和存储器 (Memory) 等, 它们之间的关系描述如图 2 所示。



2 SIG-RMS 资源信息模型的主要组成

(1) 虚拟组织 (VO)。它是组成网络的逻辑上独立的重要元素。每一个虚拟组织可以属于一个或多个不同的具体的网格应用。同样一个具体的网格应用中可以由多个不同的虚拟组织组成。一个虚拟组织内部由多个不同的计算节点组成。

(2) 计算节点 (CE)。是组成虚拟组织的逻辑上独立的重要元素, 是一个分组管理多个集群或主机信息的容器。它主要描述了该节点聚合多个集群或主机后整体性能, 如平均负载、剩余存储空间、网络带宽、可用内存大小等。一个计算节点由多个相互独立的集群或主机组成。

(3) 集群 (Cluster)。它是多个同构资源节点的逻辑集合。一般来说, 集群由多个具有相同系统配置的资源节点组成, 比如说每个资源节点具有相同型号的 CPU、存储器、网络接口、操作系统等。严格来说, 资源信息模板中关于集群节点的定义不是必需的。不过对多个同构资源节点管理来说, 如果已知一个节点的系统状况, 有了集群的定义, 我们就可以很方便地知道多个资源节点整体的系统状况。

(4) 主机 (Host)。它表示一个物理计算节点, 组成 N-GMS 资源监控系统的基本元素。通过系统静态和动态两类信息

(前者如操作系统类型、CPU 主频、存储总容量、网络地址等,后者主要包括处理器负载状况、可用存储空间、内存可用大小等)加以描述。

2.3 通信方式

SIG-RMS 监控数据的传输和管理分为局部和全局两个层次。在一个局部监控域内,主要涉及 LMA 与 Sensors 的通信以及 LMA 与 LMS 的通信。LMA 可以通过两种模式实现与 Sensors 的数据通信,即 Pull(拉模式)和 Push(推模式)。Pull 模式是指 LMA 周期地向监控域内各个 Sensor 发送获取监测数据请求,并收集 Sensor 返回的带时间戳的资源状态监测数据。Push 模式是指监控域内各 Sensor 周期地向 LMA 发送带时间戳的资源状态监测数据。LMA 负责将收集到的资源监控数据进行分析处理并且存入局部监控数据库。LMA 与 LMS 的通信主要通过接口调用来实现。在一个全局监控域内,通信主要涉及到用户、LMS 和 GMS 三个方面。LMS 首先向 GMS 注册,GMS 根据 LMS 的注册信息周期地发出监控数据查询请求,并将返回的监控数据存入全局监控数据库。用户可以通过 Web 发送查询请求从 GMS 获得监控的全局视图,同样也可以从 LMS 获得监控的局部视图。通过对监控数据的分层管理,优化网络通信性能、减少信息传输冗余,提高监控系统的整体通信性能。

3 SIG-RMS 系统实现

网格资源监控系统(SIG-RMS)当前的实现结构如图 3 所示,实现对资源全局和局部(两个不同的监控域)的两层监控。管理员或用户通过 SIG-RMS 可以监视不同监控域的资源状态。

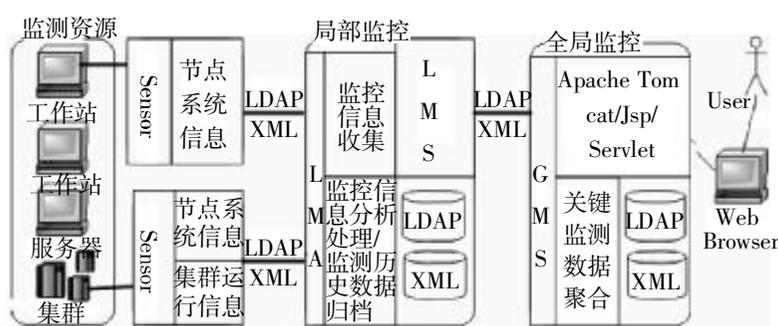


图 3 SIG-RMS 系统实现

SIG-RMS 采用 Java 实现。实现系统采用 XML 语言来描述数据、半结构化数据库来存储数据,以及轻权访问协议 LDAP 来访问数据。Sensor 将采集到的信息用 XML 进行描述,并用 LDAP 协议,采用 Push 方式将信息送给 LMA, LMA 将信息存入局部域中的数据库中,由 LMS 采用 Web 方式向本地用户提供局部视图,或服务调用方式向 GMS 或其他应用程序提供服务。GMS 将信息存入全局监控域中的数据库,构成全局视图,采用 Web 方式提供方便的数据访问。

SIG-RMS 具有如下功能特点:

(1) 监控数据多样性。根据时间戳的不同, LMA 将收集到的各种资源信息分为实时数据和历史数据分别进行分析、处理和存储。实时数据能及时了解监控域内每个集群或者各个主机节点的运行情况,满足管理员或用户对实时监测数据的请求。历史数据可以对监控域在一定时期内运行情况有一个总体的评价和分析,帮助管理员或者其他合法用户分析找出系统的瓶颈。

(2) 监控数据的分层管理。对网格资源的监控逻辑上划

分为三个层次,即虚拟组织层(VO)、计算节点层(CE)和集群/主机层(Cluster/Host)。分层组织给不同层次的管理员或其他合法用户提供了详细清晰而且针对性强的监控信息。

(3) 系统可扩展性强(Scalability)。由于 SIG-RMS 是对整个网格环境进行分层监视,而且由于整个网格环境资源可以逻辑上划分为一个个局部监控域进行管理。SIG-RMS 资源监控系统针对局部域进行监控的模块化设计就使得 SIG-RMS 具有很强的可扩展性。可以很轻松地适应网格环境的扩充和升级,包括对网格环境的虚拟组织 VO 的扩充以及对组成 VO 的计算节点 CE 的扩充。从而满足了计算能力需求不断提高,硬件性能越来越好,网格计算环境随时需要升级的要求。

(4) 系统可移植性强(Portability)。针对一些大规模的网格应用,系统可能会由相同或者不同厂商的、异构的平台,以及运行多种不同操作系统的资源构成。SIG-RMS 实现采用了 Java 组件技术,业界流行的 Web Services 服务技术,使用 XML 以及 Browser/Server 结构使得信息的获取、传输、显示标准化,从而具备了很强的跨平台的可移植性。

(5) 系统具有方便可用性。系统针对不同监控域内的管理者和用户提供不同的 Web 监控视图;提供基于 Web 的编程接口方便其他网格服务调用。

4 结论

SIG-RMS 资源监控系统是一个建立在广域环境上的分布式系统,能够对 SIG 环境中所有节点进行状态监视和监测数据的管理。其目标是为本地和异地的管理员提供一个统一的、集中的、可视化的管理工具,通过对系统的性能进行评价,为管理员或其他网格用户的决策提供参考。

本文从系统设计入手,分别阐述了系统的功能需求和系统体系结构、信息模型以及通信方面的设计思路,最后给出了一个系统原型的参考实现。下一步我们将进一步加强 SIG-RMS 的功能。

参考文献:

- [1] Tierney, R Aydt, D Gunter, et al. A Grid Monitoring Architecture [EB/OL]. <http://www.didc.lbl.gov/GGFPERF/GMA-WG/papers/GWD-GP-16-2.pdf>, 2002-01-16.
- [2] MDS. A Major Component of the Globus Toolkit [EB/OL]. <http://www.globus.org/ogsa/releases/final/docs/infosvc/MDS.html>, 2004.
- [3] R-GMA: Relational Grid Monitoring Architecture [EB/OL]. <http://www.r-gma.org>, 2003.
- [4] Network Weather Service [EB/OL]. <http://nws.cs.ucsb.edu/>, 2002-07-07.
- [5] R Wolski, N Spring, J Hayes. The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing. Future Generation Computing Systems [EB/OL]. <http://nws.npaci.edu/1999>.

作者简介:

田鸣华(1977-),男,硕士,主要研究方向为网格计算、软件工程;肖依,男,教授,博士,主要研究方向为网格计算、并行与分布计算技术;刘波,男,教授,博士,主要研究方向为并行与分布式计算、生物网格;黄斌,男,硕士,主要研究方向为数据网格、分布式存储。