

地震勘探集成环境资源监控研究与实现^①

钟 敏^{1,2}, 李 杰², 梁 鸿², 仝兆岐^{1,2}

¹(中国石油大学(华东) 地球科学与技术学院, 青岛 266580)

²(中国石油大学(华东) 计算机与通信工程学院, 青岛 266580)

摘 要: 利用网格技术建设地震勘探集成环境高性能计算平台基础设施的过程中, 监控系统不仅要为位于不同管理域、分布异构的高性能资源管理提供资源全局视图, 方便对资源管理和控制, 还要为分布环境下资源的选择、任务的调度与执行提供指导信息. 本文提出了面向地震勘探领域的高性能应用网格监控系统体系结构, 建立了监控信息数据模型, 在 Globus Toolkit 信息服务组件的基础上扩展了信息采集功能和信息查询接口, 并基于 Websphere 为高性能计算资源的信息查询与监控提供了资源实体的统一视图和可视化管理 Portal. 经测试, 系统运行稳定, 可扩展性强.

关键词: 网格; 资源监控; 信息采集; 查询接口; Portal

Monitoring System for Seismic Exploration Integrated Environment

ZHONG Min^{1,2}, LI Jie², LIANG Hong², TONG Zhao-Qi^{1,2}

¹(School of Geosciences, China University of Petroleum, Qingdao 266580, China)

²(College of Computer and Communication Engineering, China University of Petroleum, Qingdao 266580, China)

Abstract: In the construction of high-performance computing grid infrastructure for seismic exploration integrated environment, the monitoring system should not only provide a global view for distributed and heterogeneous high-performance resources in different administrative domains, but also provide guiding information for resources selection and task scheduling. A high-performance grid monitoring architecture for seismic exploration integrated environment is proposed. Then the key techniques, including monitoring information data model, information provider, information query interface and Portal, are discussed. MDS information service component of Globus Toolkit is extended and Portal system based on Websphere is developed. System testing shows that the monitor system is stable, good scalability.

Key words: grid; resources monitoring; information provider; query interface; Portal

在石油勘探行业, 随着以万道地震技术和特殊地震技术为代表的新一代地震勘探技术的发展与应用, 地震资料处理软件进一步向集成化和并行化大规模数据处理发展^[1]. 高性能计算平台支撑技术的发展与应用成为推动油气勘探数据大规模并行处理的重要动力. 利用网格技术^[2]可以整合行业内部已有各种高性能计算资源, 实现资源的统一管理, 提高资源利用率, 为地震勘探集成环境建立高性能计算平台基础设施. 在应用网格基础设施的建设中, 监控系统是其重要组成

部分. 监控系统不仅要为平台管理提供资源全局视图, 方便对资源管理和控制; 还要为分布环境下资源的选择、任务的调度与执行提供指导信息.

Globus Toolkit4(GT4)^[3]的信息服务组件 MDS4 用于在网格环境下实现网格资源的发现、状态查询及监控, 为用户提供共享资源的全局视图和命令行信息查询接口, 具有良好的扩展性和容错性^[4]. 但是, 在实际应用中还不能满足地震勘探集成环境的需求. 一方面, 分布异构环境下需要根据应用需求调整需要监控的资

^① 基金项目: 中国教育科研网格 Chinagrid 资助(M1307009A); 中央高校基本科研业务费专项资金(R1107009A)

收稿时间: 2013-06-25; 收到修改稿时间: 2013-07-22

源节点的状态信息,这就需要基于 Gloubs 规范进一步扩展 MDS 的功能. 另一方面,集成环境下的元调度系统需要实时获取可用计算资源列表及其详细状信息,并根据这些信息指导调度策略的实施. 而对于网络用户来说需要形象直观的、方便的监控操作界面. 这就需要针对 MDS 组件设计专门的编程接口和资源监控界面.

本文提出了应用网络监控系统的体系结构,分析了系统关键技术,建立了监控信息数据模型,扩展了 MDS 信息采集工具,将扩充的资源信息补充到 MDS 信息服务中;设计并实现了基于 MDS 查询接口,对资源信息进行了解析,为元调度器和用户 Portal 提供访问接口;最后基于 Websphere 为用户提供了资源监控与查询的可视化访问 Portal.

1 应用网络监控体系结构

应用网络环境中的资源具有分布、异构、实时动态、自治等特性. 共享的物理资源根据应用需求,隐藏了与应用无关的物理特性,经过抽象和虚拟化加入到网络虚拟社区中,为用户提供特定服务. 资源请求者需要实时获取这些可用资源状态信息来指导任务的调度与执行.

应用网络监控系统在信息请求者和资源提供者之间架起了一座桥梁,负责对网络资源实体的可用性以及相关特征描述等信息进行维护和监控,为网络用户提供资源实体的统一视图和动态管理界面,能实时的反映应用网络环境的动态变化. 地震勘探集成网络应用平台主要用于聚合各类高性能计算资源为行业应用提供计算服务,资源请求者关心的主要是计算资源的相关状态信息和系统性能信息,设计系统体系结构如图 1 所示.

监控中间件为资源信息的获取和管理提供基本操作,主要包含信息注册、更新、汇聚、注销、查询等基本功能.

信息采集模块使用信息采集程序(Provider)对集成平台下异构计算资源的信息进行采集并推送到监控系统,为元调度系统和用户查询提供支持.

资源监控 Portal 是平台系统的重要组成部分,是用户与平台的桥梁,为用户提供了简单、方便的图形操作界面. 通过 Portal,用户可以直观的获取图形化的资源信息,同时也能够查询资源的历史状态信息.

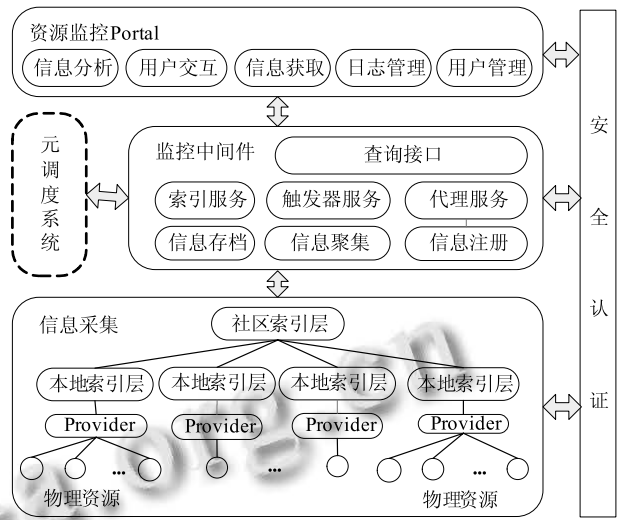


图 1 系统体系结构

安全认证模块基于 GT 网络安全基础设施(GSI)提供认证、授权和安全通信,为用户安全、有效的访问位于不同管理域、异构的资源提供了保障.

2 关键技术

2.1 监控信息数据模型

地震勘探集成环境中,资源组织为如图 1 信息采集模块所示的层次模型. 第一层为社区索引层,汇聚了集成环境中的所有工作节点信息;第二层为本地索引层,汇聚本地集群环境中工作节点的信息. 因此,将数据模型组织为如图 2 所示的资源信息树.

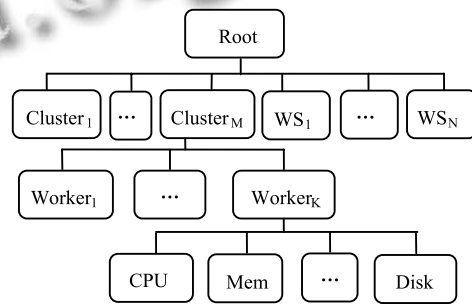


图 2 资源信息树

2.2 信息采集

在分布、异构的集成环境中,信息采集模块要对位于不同层次的节点资源的信息进行收集,包括静态信息和动态信息等,详细的属性信息如表 1 所示.

在本地索引层,集群系统内的工作节点通过集群

内部监控系统将信息汇聚到本地索引层。在社区索引层, MDS 提供了信息聚集接口, 利用 MDS 可以实现网格中信息的收集和发布, 这些信息是由信息 Provider 提供的。集群系统或者工作站服务器通过自定义 Provider, 并向 MDS 注册, 定时将信息推送到 MDS 信息服务组件, MDS 再将信息汇聚到全局索引服务中。

表 1 资源主要属性信息

静态信息	IP 地址信息
	处理器相关信息
	内存容量
	缓存容量
	操作系统信息
动态信息	系统启动时间
	处理器负载
	内存消耗情况
	网络负载
	文件系统状态

另一方面, 由于 MDS 没有提供信息存档服务, 而集成平台不仅要为用户提供资源的历史信息, 还要保证系统调度性能趋于最优。所以, Provider 在推送信息的同时还需要将资源信息存档, 以满足用户的监控, 并为元调度系统提供资源的历史状态信息指导资源的调度与分配。

2.3 信息查询接口

虽然 MDS 为监控系统提供了命令行工具, 用户可以使用相应的命令来搜索和查询集成环境中的资源状态信息。但是结果数据由结构复杂的文本内容组成, 信息被隐含在大量无关信息中, 用户难以直观的掌握资源的状态。另一方面, 元调度器也要按需从监控系统获取资源状态信息, 为作业匹配最优资源。因此, 信息查询接口需要对 MDS 查询结果进行解析, 为图形监控界面和元调度器提供编程接口。

系统流程如图 3 所示。信息查询接口从元调度系统和用户查询界面接受查询指令, 启动信息查询与分析线程, 查询 MDS 服务, 并将返回结果解析后反馈给信息请求方。

2.4 用户 Portal

用户 Portal 屏蔽了网格系统的复杂性, 为用户提供基于 Web 的访问界面。通过 Portal, 用户可以方便地进行身份验证, 按需搜索网格资源, 透明地监控网格资源。用户对资源信息的访问主要分两类: 搜索满足条件的资源和监控某个资源的历史状态信息。Portal

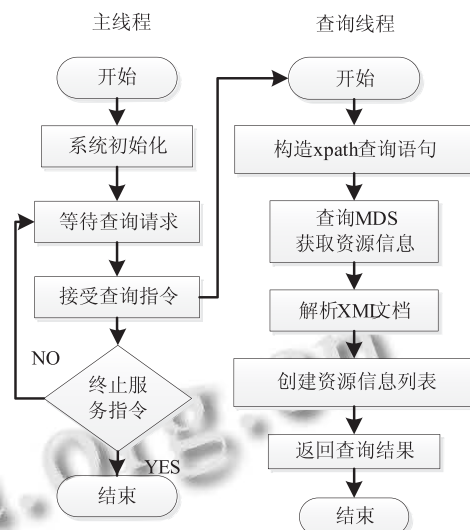


图 3 信息查询接口流程

功能由以下三个模块构成:

交互模块用于接受用户的输入和请求, 将相应的信息传递给其它模块进行进一步的处理。

信息获取模块通过认证首先与社区索引服务主机建立连接, 调用信息查询接口, 获取满足条件的资源信息并将信息交付信息分析模块格式化处理。

信息呈现模块将信息获取模块的结果解析后在 Web 页面中可视化显示; 针对历史信息监控请求, 将历史数据以图形形式呈现在 Web 页面。

用户通过 Portal 监控集成环境资源流程如图 4 所示。

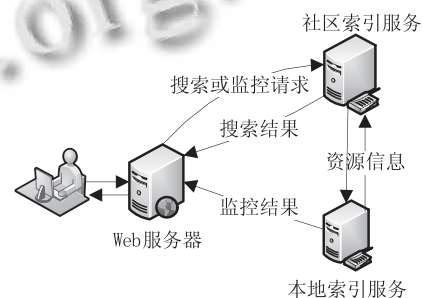


图 4 用户访问流程

3 系统实现

3.1 资源信息的表示

为了在集成环境内传递和共享资源信息, 监控信息的表示要考虑以下几个因素: 1)具备平台无关性, 可以在不同平台和应用程序之间交换数据; 2)有利于快速检索和查询; 3)便于在网路中传输。XML 文档所具

有的特点正好满足这些需求, 所以采用 XML 文档来定义和描述资源信息. 以工作站节点信息为例, 信息格式定义如下:

```
<Host Name=".." iqueID=" ">
  <Processor CacheSize="" ClockSpeed="" CpuMHz="" TotalCPUs="" modelName="" />
  <MainMemory buffers="" Cached="" MemFree="" MemTotal="" MemUsed="" SwapFree="" SwapTotal="" SwapUsed="" />
  <OperatingSystem Machine="" Name="" Platform="" Release="" />
  <FileSystem>
    <DEVICE1 MountPoint="" Total="" free="" used="" usedPercent="" />
    <DEVICE2 MountPoint="" Total="" free="" used="" usedPercent="" />
    <DEVICEn MountPoint="" Total="" free="" used="" usedPercent="" />
  </FileSystem>
  <NetworkAdapter IPAddress="" MTU="" ReceiveByte="" ReceivePK="" TransportByte="" TransportPK="" />
  <ProcessorLoad BootTime="" Last15Min="" Last1Min="" Last5Min="" />
</Host>
```

3.2 信息 Provider

地震勘探集成环境中主要有两类信息 Provider.

3.2.1 工作站节点 Provider

(1) 信息收集与存档

linux 系统目录/proc 存储了当前主机的硬件信息和系统信息. 该目录下的文件不占硬盘空间, 只存在内存中. 编写 shell 脚本 Provider, 读取/proc 目录下系统信息, 并对其进行处理和格式化. 同时基于“心跳机制”将历史数据写入 rrdtool 数据库, 为用户可视化查询与信息分析提供支持.

(2) 为索引服务配置 Provider 信息

修改 MDS 中索引服务的 server-config.wsdd 文件, 在文件中添加以下内容, 为 DefaultIndexService 添加 Provider:

```
<parameter name="rpProviderConfigFile"
value="etc/globus_wsrf_mds_usefulrp/gluece-rpro
```

```
vider-sample-config.xml"/>
```

然后在文件 gluece-rpprovider-sample-config.xml 中设置信息收集脚本服务路径, 完成 Provider 的配置.

3.2.2 集群系统 Provider

(1) 集群系统信息收集

在集成环境中, 集群系统的集群管理节点安装了本地索引服务, 接入应用网格平台. 集群管理节点要为索引服务收集集群内所有节点相关信息.

目前, 集群环境下监控软件主要有 Hawkeye、Ganglia、Nagios、Performance Benchmark 等. 其中 Ganglia 具有跨平台、可扩展、搜集数据精确等特点. 因此, 本系统选择 Ganglia 作为集群系统的信息 Provider, 监控集群内部所有主机节点的 CPU、内存、文件系统、网络设备等相关信息.

(2) 启用 Usefulrp 框架

MDS 采用 Usefulrp 框架支持 Ganglia 等第三方软件, 将其搜集的信息汇聚到索引服务中. 配置方法如下: 修改 MDS 中 IndexService 的 server-config.wsdd 文件, 增加以下内容, 启用 Usefulrp 框架

```
<parameter name="rpProviderConfigFile"
value="etc/globus_wsrf_mds_usefulrp/gluece-rpprovider-sample-config.xml"/>
```

(3) 为 Usefulrp 框架配置 Provider 信息

\$GLOBUS_LOCATION/etc/globus_wsrf_mds_usefulrp 目录下, 使用 mds-gluerp-configure 命令配置 Provider 信息. 如 mds-gluerp-configure fork ganglia gluece-rpprovider-sample-config.xml 命令指定集群信息 Provider 为 ganglia.

3.3 信息查询接口

信息查询接口定义了四个服务类: Im_Mds 类, ServiceIM 类, QueryClient 类, QueryParser 类.

其中 Im_Mds 类用于接收查询参数, 包括端口号, 主机列表与社区 MDS 节点等, 在完成初始化工作后, 等待查询命令, 启动查询线程, 将命令消息传递给 ServiceIM 实例对象. ServiceIM 类在接收命令消息后, 分别调用 discover() 和 monitor() 函数, 这两个函数使用 QueryClient 实例对象, 与社区 MDS 节点或者具体主机建立连接, 获取资源信息, 信息以 xml 文档格式返回. QueryClient 类实例使用 QueryParser 类实例对象对查询结果 xml 文档进行解析, 创建信息列表.

用户 Portal

系统基于 Websphere Portal 实现了用户访问门户。系统实现分系统服务类、Portlet 类和 jsp 界面三个层次实现。

系统服务类是系统实现的核心类, 实现了集成环境中的资源信息查询和监控。其中最重要的是 Discover 类和 Monitor 类。Discover 类生成整个社区资源索引列表, 根据用户请求查询资源信息; Monitor 类根据用户请求获取目标资源历史数据。

所有 Portlet 类继承 ActionPortlet 类, 用于与系统服务类交互和管理 Web 页面显示, 是沟通系统服务类和 jsp 页面的桥梁。其中 UserPortlet 读取并显示用户信息和角色权限、管理用户, 并管理 userinfo 和 help 页面的显示。ResMon 类根据用户请求管理资源信息列表, 绘制并更新资源历史信息图表。

jsp 界面为用户提供查询监控访问界面。

4 结论

本文工作是地震勘探应用网格平台建设的重要组成部分。系统在网格基础设施 Globus Toolkit 信息服务组件的基础上扩展了信息采集和信息查询接口, 并为地震勘探应用环境下高性能计算资源的信息查询与监控提供了资源实体的统一视图和可视化管理 Portal。

系统实现了位于不同管理域、分布异构的高性能计算资源的信息查询与监控。系统通过测试, 运行稳定、可扩展性强。不足之处在于对元调度系统任务调度的决策支持有待进一步扩展。下一步工作拟在监控资源历史信息的基础上, 对历史数据进行分析以支持作业调度策略的执行。例如通过分析某个资源节点的 cpu 历史负载信息, 指导元调度系统在调度作业时避免将作业分配到过载节点执行。

参考文献

- 1 赵连功, 刘洪. 地震勘探数据资料处理软件集成化研究现状和发展趋势. 地球物理学进展, 2003, 18(4): 598-601.
- 2 Foster I, Kesselman C, Tuecke S. The anatomy of the Grid: Enabling scalable virtual organizations. International Journal of Supercomputer Applications, 2001, 15(3): 200-222.
- 3 Foster I. Globus toolkit version 4: Software for service-oriented systems. Journal of Computer Science and Technology, 2006, 21(4): 513-520.
- 4 Schopf JM, Pearlman L, Miller N, Kesselman C, Foster I, D'Arcy M, Chervenak A. Monitoring the grid with the Globus Toolkit MDS4. Journal of Physics, 2006, 46(1): 521-525.
- 5 Zealand. 2003. 18-23.
- 6 Dillencourt MB, Samet H, Tamminen M. A general approach to connected component labeling for arbitrary image representations. Journal of the Association for Computing Machinery(ACM), 1992, 39(2): 253-280.
- 7 Chang F, Chen CJ, Lu C J. A lineartime componentlabeling algorithm using contour tracing. Technique Computer Vision and Image Understanding, 2004, 93(2): 206-220.
- 8 张桂林. 基于跑长码的连通区域标记算法. 华中理工大学学报, 1994, 22(5): 11-14.
- 9 蔡世界, 于强. 基于游程编码的连通区域标记算法优化及应用. 计算机应用, 2008, 28(12): 3150-3153.
- 10 徐利华, 陈早生. 二值图像中的游程编码区域标记. 光电工程, 2004, 31(6): 63-65.
- 11 高红波, 王卫星. 一种二值图像连通区域标记的新算法. 计算机应用, 2007, 27(11): 2776-2778.
- 12 He L, Chao YY, Suzukik. A runbased twoscan labeling algorithm. IEEE Trans. on Image Processing, 2008, 17(5): 749-756.
- 13 聂欢欢, 伊磊, 刘任平. 基于区域生长法提取二值图像中的连通区域. 计算机时代, 2012, 6: 23-24.
- 14 胡涛, 郭宝平, 郭轩等. 一种串行/并行两用的区域标记算法. 计算机工程, 2010, 36(9): 17-22.

(上接第 159 页)