

网格服务复制机制的研究^{*})

宋宝燕¹ 李旭峰¹ 于洋¹ 毛艳英¹ 申德荣²

(辽宁大学信息科学与技术学院 沈阳 110036)¹ (东北大学信息科学与工程学院 沈阳 110004)²

摘要 网格的 OGSA 模型要求网格上一切资源都是服务,实现对网格服务的复制可以解决网格中网格服务的访问瓶颈、负载均衡等问题。现有的网格副本服务技术只能对数据、文件资源等静态资源进行复制,不能对包含软硬件需求的有状态的网格服务进行复制。针对这一问题,我们提出了一种有状态网格服务的复制方法,该方法基于虚拟工作空间、Snapshot 和有状态资源模拟技术的原理,实现了对网格服务代码的复制、对网格服务运行所需的软硬件环境的复制和搭建、对网格服务内存数据的复制和对网格服务使用的有状态资源的复制。

关键词 网格服务,SGSR,虚拟工作空间,有状态资源

Research of Service Replication Mechanism in Grid

SONG Bao-yan¹ LI Xu-feng¹ YU Yang¹ MAO Yan-ying¹ Shen De-rong²

(Department of Information Science and Technology Institute, Liaoning University, Shenyang 110036, China)¹

(School of Information Science and Engineering, Northeastern University, Shenyang 110004, China)²

Abstract The OGSA model on Grid demands all the resources are services. Realizing the replication of Grid Service can resolve the questions such as accessing bottleneck, load balance. Existing Grid Replication service can replicate static data and file resource, but it is useless to stateful Grid services which include software, hardware requirement. Aiming at this problem, we advise a kind of replication methods about stateful Grid Service, which implements the replication of software and hardware environment using customized trait and well executing separation attribute of Virtual Workshop, accomplishes the replication of Grid service memory state based on the Snapshot, realizes the replication of stateful resource adopting modeling technology of Web service.

Keywords Grid service, SGSR, Virtual workshop, Stateful resource

1 引言

随着网格技术的发展,网格框架逐步采用了新的 OGSA 模型。OGSA 模型是一种完全面向服务的模型,由基本服务和高级服务组成。大部分网格服务与静态资源不同,是有状态的,能够随时间和网格用户的调用不断地改变自身状态。为了能够在 OGSA 模型下消除网格服务的访问瓶颈,能够对网格服务的就近访问,使网格服务负载均衡,我们必须实现对有状态的网格服务的复制。

经过专家学者们的大量研究,现有的副本服务技术已经能够实现对文件、对象、海量数据和广域数据的复制。这些数据都是静态的数据资源(数据状态不发生变化,无状态资源),因而副本服务技术无法对网格服务的状态进行复制。

网格服务的复制需要对许多不同的动态改变的资源进行复制,即对网格服务程序进行复制、对网格服务运行所需的软硬件环境进行复制和搭建、对网格服务使用的有状态资源进行复制,对网格服务的内存数据进行复制。

基于以上问题,本文作者深入分析了网格服务运行的具体环境和实现机理,充分考虑了网格服务的动态性及有状态资源可重构等特点,提出了一种基于虚拟工作空间的网格服务复制机制 SGSR(Stateful Grid Service Replication)。SGSR

利用虚拟工作空间可定制化的特性和良好的执行分离属性,解决了对网格服务运行所需的软硬件环境的复制;根据 Snapshot 的原理,完成了网格服务内存状态的复制;采用 Web 服务的有状态资源模拟技术,实现了有状态资源的复制。

2 相关技术

2.1 虚拟工作空间技术

虚拟工作空间^[1,2](Virtual Workshop, VW)是一种典型的虚拟化技术,能够按照用户的需求构造可定制的应用环境。类似的技术有虚拟机工厂^[3]。虚拟机工厂的实现不利于所需环境的灵活配置,所以本文采用了虚拟工作空间技术来解决网格服务所需应用环境的复制。

虚拟工作空间采用 XML Schema 进行定义,定义包含默认的工作空间配置、软件模块信息、用户配置信息及网络接口信息等。虚拟工作空间的实现技术有很多种,如动态分配和配置用户账户技术、预配置虚拟机技术等,本文采用虚拟机技术。

虚拟工作空间的虚拟机技术可以看成是一组虚拟机监控程序(VMM)的集合。VMM 能够截取和模拟特权指令,创建和管理多个虚拟机,还提供接口,允许网格用户开始、暂停、停

^{*} 基金项目:国家自然科学基金资助项目(60673139)。宋宝燕 博士,教授,主要研究方向为网格技术、数据流技术、数据库技术;李旭峰 硕士研究生,主要研究方向为网格技术;于洋 硕士研究生,主要研究方向为数据流挖掘;毛艳英 硕士研究生,主要研究方向为网格技术;申德荣 博士,教授,主要研究方向为网格技术、分布数据库技术。

止和恢复虚拟机。通过 VMM,虚拟机的状态可以被序列化成为虚拟机映像。虚拟机映像中包含有虚拟 RAM、虚拟硬盘、配置文件的映像。

虚拟工作空间由虚拟工作空间信息服务、虚拟工作空间配置服务和虚拟工作空间部署服务组成。虚拟工作空间信息服务包含已经配置好的、可用的虚拟工作空间的描述,使用户能够按需查找所需的虚拟工作空间。虚拟工作空间配置服务主要是将用户需求转换成 XML 文档,按照文档在服务中进行配置,构造出用户所需要的、形式化的虚拟工作空间。虚拟工作空间部署服务是将形式化的虚拟工作空间发布、绑定到具体的物理资源上,对虚拟工作空间进行实例化。

2.2 副本服务技术

副本服务技术^[5]实现了对静态资源的复制、副本的注册和副本的选择功能,包括副本管理服务、副本目录服务和副本选择服务。副本管理服务可分为复制服务和 GridFTP 传输服务,静态数据资源的复制和传输都是在副本管理服务中实现的,它还负责对其他子服务的调度;副本目录服务记录每个资源有多少个副本、副本的位置信息等;副本选择服务对查找到的所有副本进行选择,挑选一个最佳的副本。

3 SGSR 的体系结构

SGSR 机制主要针对网格服务及其状态进行复制,网格服务中的无状态资源的复制由副本服务技术来解决。本文所说的网格服务的状态主要指网格服务所需的硬件环境、网络资源、软件环境、有状态资源^[4]、应用程序的状态及服务内存状态等。

SGSR 的体系结构如图 1 所示。SGSR 完全符合 OGSA 模型,因而它的体系结构是一个包含多个子服务的复合服务。SGSR 能够将单机环境下的网格服务复制到单机或广域环境下的一个虚拟环境,也可将虚拟环境下的网格服务复制到另一个虚拟环境下。下面对各个子服务的功能进行介绍。

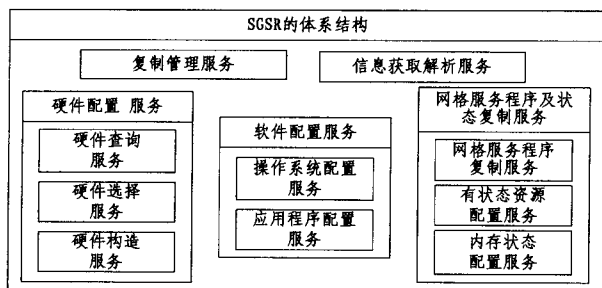


图 1 SGSR 的体系结构图

复制管理服务实质上是一个调度引擎,负责在各子服务之间进行调度,协调各子服务的运行来完成复制任务。复制管理服务以网格服务的全局标识符作为输入,开启一个事务,然后依次调度信息,获取解析服务、硬件配置服务、软件配置服务、网格服务程序及状态复制服务。如果有一个子服务的复制操作是不成功的,则事务进行回滚。复制管理服务还负责安全方面的一些处理,负责同主机或客户进行谈判,以获取更多所需的资源。安全方面的具体实现不在本文中详细讨论。在网格服务复制完成后,还要进行注册。SGSR 的注册机制和副本服务技术的注册机制大致相同,可以使用副本服务技术的注册服务进行注册。复制管理服务需要将网格服务的副本作为一个新服务进行注册,便于用户就近访问,消除访

问瓶颈,同时要在副本服务中注册,使其成为原网格服务的一个副本来解决负载均衡的问题。

信息获取解析服务主要用来获取原网格服务注册的软件和网络信息等,为软硬件配置服务提供信息支持。信息获取解析服务通过自动查询目录服务的 LDAP 树,并通过与原网格服务进行通信,获取指定网格服务运行所需的内存大小、硬盘大小、网络带宽这样的硬件信息和操作系统类型与版本、应用程序的类型与版本这样的软件信息,并将这些信息封装成 XML 文档,供其他的子服务进行读取。

硬件配置服务是用来搭建和配置网格服务所需的硬件环境的,为网格服务提供有效的支撑运行环境,是软件配置服务的基础。硬件配置服务需要从现有的硬件资源中查找、获取和改造硬件,并不能真正地对硬件进行复制,也无法实现硬件复制。硬件配置服务主要是通过调用虚拟工作空间的子服务来完成硬件的复制的,虚拟出需要的硬件环境。硬件配置服务在没有查找合适虚拟工作空间的时候,可以通过复制管理服务和具有所需硬件能力的机器进行谈判。如果谈判成功,则在谈判的主机上安装 VMM 程序。并创建一个满足所需硬件需求的虚拟机,以便提供硬件服务。

软件配置服务是建立在配制好的硬件之上的,为网格服务配置所需的软件环境。软件环境包括网格服务自身所需的软件环境和有状态资源所需的软件环境。网格服务自身需要的软件环境主要是操作系统及网格服务运行所需的容器,有状态资源需要的软件环境是其存在和状态改变所依托的软件。软件配置服务通过使用虚拟机磁盘映像来达到目的,将广域环境的多台主机上的磁盘映像文件进行复制,在新配置硬件的虚拟机上进行恢复。

网格服务程序及状态复制服务实现了网格服务程序、内存状态和网格服务用到的有状态资源的复制。该服务由网格服务程序复制服务、有状态资源配置服务和内存状态配置服务 3 个子服务组成。网格服务程序一般为编译好的一些类文件,属于静态资源,利用副本服务技术进行复制。有状态资源配置服务主要是针对有状态资源(如文件系统中的文件,关系数据库中的表的一条记录,或是 J2EE 中的 EJB 对象等)进行复制。有状态资源可以是上述一组有状态资源的集合体。有状态资源配置服务通过与原网格服务进行信息交换,利用资源属性文档将有状态资源及其状态复制到已经复制好的软件环境中,并在虚拟机上进行恢复。内存状态配置服务通过 Snapshot 将某一时刻的内存数据映射到硬盘进行存储,复制后在新的网格服务副本上将数据恢复成内存数据。

4 SGSR 对网格服务的复制

SGSR 复制机制对网格服务的复制主要是通过对网格服务程序及它的各种状态的复制来实现的。SGSR 侧重于对网格服务所需的硬件环境、软件环境、有状态资源及网格服务内存状态等进行复制。

4.1 硬件的复制

硬件环境的复制由硬件配置服务来实现,这是 SGSR 机制对网格服务其他状态进行复制的基础,复制任务由硬件查询服务、硬件选择任务和硬件构造服务 3 个子服务共同完成。下面对硬件环境的复制过程进行具体说明。

(1)复制管理服务调用信息获取解析服务,获得原网格服务的注册信息。

(2)由硬件查询服务读取信息,获取解析服务的 XML 文

档中关于原网格服务硬件和网络信息。

(3) 硬件查询服务在 VW 信息服务中查找符合要求的 VW。如果存在这样的 VW, 调用硬件选择服务, 选择最优的 VW, 返回 VW 的虚拟机句柄, 结束硬件复制。

(4) 如果不存在这样的 VW, 使用 VW 配置服务按照 XML 中的信息构造形式化的 VW。如果能够成功配置返回 VW 虚拟机句柄, 结束硬件复制。

(5) 配置不成功, 使用硬件查询服务在目录服务中查找, 同时提供内存和网络服务并符合要求的主机。如果找不到, 调用硬件构造服务转(7)。

(6) 如果找到, 调用硬件选择服务, 选出最佳的主机并返回主机的地址。在目录服务中查找满足要求的磁盘资源, 找到返回硬盘的映像, 结束复制。找不到, 调用硬件构造服务。

(7) 硬件构造服务按照预先定义的策略同满足需求但不能提供相应服务的主机进行谈判, 获取权限。谈判成功, 再调用 VW 部署服务进行构造, 不成功报告失败。

下面对上述的调用过程做一些补充说明。

硬件查询服务查询的过程是一个与查询, 只有当内存、硬盘、网络等所需的所有硬件信息都满足的时候, 才算找到一个真正可用的资源。

硬件选择服务选择资源的标准如下: 可扩展空间大的 VW 或资源优先选择, 物理位置与原服务所在主机距离近的 VW 或资源优先选择, 建立在虚拟机之上的资源优先选择。

硬件构造服务在获得足够的权限后, 先在目标主机上安装 VMM 进行虚拟化, 再调用 VW 部署服务。

4.2 软件环境的复制

软件环境的复制由软件配置服务实现, 它是建立在配置好的硬件和网络环境之上的。本文将网格服务看成是由无状态的网格服务程序和有状态资源组成的, 每一个有状态资源实际上就是网格服务一部分状态。有状态资源不同, 服务的状态就不同。

软件配置服务由操作系统配置服务和应用程序配置服务组成, 分别完成对操作系统和所需应用环境的复制。软件环境的复制过程如下:

(1) 操作系统配置服务获取信息, 获取解析服务 XML 信息文档的内容。

(2) 操作系统配置服务检查硬件上是否装有虚拟机。如果不存在, 则安装虚拟机。

(3) 检查虚拟机上的操作系统的类型与 XML 文档的要求是否相同。如果不同, 就加载操作系统配置服务中预先配置的该类型 OS 的虚拟机映像, 替换现有的虚拟机。

(4) 如不存在这样的预配置映像, 在目录服务中查找所需系统的安装文件包资源。查到后, 在虚拟机上进行安装, 完成后加入预配置系统映像库。查不到报告, 失败回滚。

(5) 系统配好后, 操作系统配置服务读取虚拟机所在主机的硬件(包括网络)驱动程序, 并复制到新的虚拟机中(VMware 中是通过安装 VMware Tools 来实现)。

(6) 应用程序配置服务从信息获取解析服务中查找原网格服务运行所需的容器的类型。

(7) 在配有操作系统的虚拟机上映射或安装对应的服务容器。不成功, 则回滚结束。

(8) 应用程序配置服务在全局元素声明(GED)中查找原网格服务的资源属性文档, 获得所有的有状态资源的资源属性文档中的有用信息, 并存储在一个 XML 文档中。

(9) 按照文档中有状态资源的类型确定有状态资源所需的应用软件。

(10) 应用程序配置服务在目录服务中查找是否有提供这些应用程序服务的资源映像。有则进行映射; 没有, 就查找安装包, 在虚拟机上进行安装。不成功就回滚。

4.3 网格服务及有状态资源的复制

网格服务程序是指能够完成网格服务功能的类文件及其它程序文件, 这些文件都是静态资源, 是无状态的, 它的复制使用副本服务技术来完成。网格服务的有状态资源和内存状态的复制由内存状态配置服务和有状态资源配置服务来完成。

有状态资源用 XML 模式来描述, 每一个有状态资源对应一个资源属性文档。资源属性文档记录该有状态资源的属性、取值、类型等。

资源属性文档放在一个 XML 全局元素声明(GED)中, GED 定义了资源属性文档的根元素。任何有状态资源取值的改变都会反映在资源属性文档上, 资源属性文档相当于有状态资源的一个视图。有状态资源的调用图如图 2 所示。

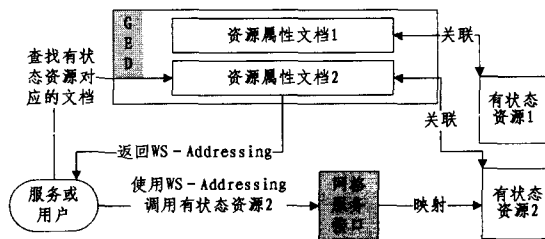


图 2 有状态资源调用图

有状态资源的改变是通过网格服务间的信息交换来完成的, 这些 SOAP 消息中包含有对资源进行改变的标准化操作(如添加、删除、更新等)。有状态资源的复制如下所述:

(1) 有状态资源配置服务在 GED 中查找所有属于原网格服务的资源属性文档。

(2) 通过查找获得有状态资源的一个服务终端引用, 引用中包含有状态资源的标识符。

(3) 有状态资源配置服务通过网格服务接口使用服务终端引用定位到有状态资源。

(4) 有状态资源配置服务与原服务进行信息交换, 将有状态资源的数据提取成一个 XML 文件。

(5) 使用有状态资源工厂在新环境上按照提取的 XML 文件的内容创建有状态资源。

(6) 有状态资源配置服务创建新资源的资源属性文档, 并将文档与资源相关联, 然后将文档存放在 GED 中。

(7) 将所有的有状态资源都如上创建。不成功则回滚。

经过软硬件和有状态资源的复制, 网格服务基本复制完成, 但新网格服务不能达到原服务当时的状态。如果按照操作步骤重新来做, 往往需要几小时、几天甚至上月的时间。如果能够对内存中的堆栈、计数器中的数据也进行复制, 就能使网格服务恢复当时的运行状态。

对内存的复制由内存状态配置服务来完成, 内存配置服务通过借鉴 Snapshot 原理来完成这一任务。本文依旧将其称为 Snapshot。Snapshot 实际上是内存数据的指针集合, 它在磁盘控制器的 cache 中建立一个数据区域指针, 与内存中的块一一对应。对系统使用 Snapshot 功能时, Snapshot 在磁

(下转第 113 页)

参考文献

- [1] Venkateswaran R. Virtual private networks. Potentials, 2001, 20(1):11-15
- [2] Rosenbaum G, Lau W, Jha S. An analysis of virtual private network solutions. Local Computer Networks, 2003, 10:395-404
- [3] Yurcik W, Doss D. A planning framework for implementing virtual private networks. IT Professional, 2001, 3(3):41-44
- [4] Gleeson B. A Framework for IP Based Virtual Private Networks. IETF RFC2764, 2000, 2. [http://www.ietf.org/rfc/](http://www.ietf.org/rfc/rfc2764.txt)

- rfc2764.txt
- [5] 杨波. IP VPN 技术应用的研究. 长沙通信职业技术学院学报, 2006, 5(1):40-44
- [6] 杨彦彬. 可运营的 IP-VPN 业务. 广东邮电职业技术学院学报, 2005, 1(1):21-23
- [7] Rosenbaum G, Lau W, Jha S. Recent directions in virtual private network solutions. Networks, 2003(9/10):217-223
- [8] Tamasi L, Orincsay D, Gabor B, et al. Design of survivable VPN based VoIP networks. Design of Reliable Communication Networks, 2005, 10:8

(上接第 102 页)

盘上开辟一个存储区,按 Snapshot 的指针将内存中的数据复制到磁盘上,然后更改指针指向磁盘上的块。读取时按照数据区域指针对应的块读进相应的内存块。Snapshot 的原理图如图 3 所示。

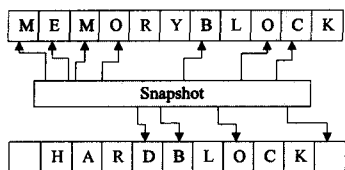


图 3 Snapshot 的原理

内存状态的复制过程如下:

- (1) 内存配置服务通过 SOAP 消息与原网格服务进行通信,消息中包含进行 Snapshot 标准操作的指令。
- (2) 对 Snapshot 生成的多个文件进行复制。
- (3) 在新环境中对 Snapshot 在内存中进行恢复。

5 实验

SGSR 复制方法能够很好地完成网格服务的复制任务,复制后的网格服务能够顺利地运行。我们对同一个网格服务在未使用和使用 SGSR 的情况下的用户访问响应时间和吞吐量分别进行了测量,实验结果如图 4、图 5 所示。

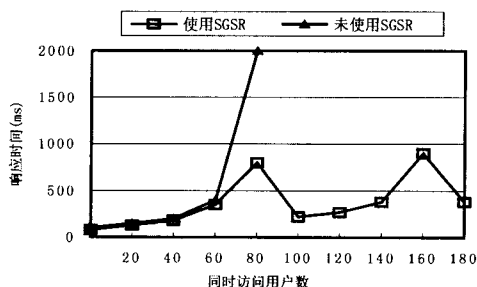


图 4 用户访问的响应时间

在图 4 中,未使用 SGSR 的情况下,在同时访问网格服务的用户数小于 60 的情况下响应时间较少。当用户数达到 80 后,服务处于超负荷状态,响应时间迅速增加。而使用 SGSR 后,当访问用户数接近 80 时就开始对网格服务进行复制,复制成功后一部分用户转向访问网格服务的副本。由于网格服务的用户访问数减少,响应时间迅速降低,因此 SGSR 可以很好地解决访问瓶颈的问题。

图 5 是使用 SGSR 和未使用 SGSR 时吞吐量的比较。未

使用 SGSR 时,网格服务副本只有一个。随着任务不断增加,吞吐量增大(未过载),最终达到饱和。使用 SGSR 后,当吞吐量趋近饱和的时候,可以通过复制出新的网格服务副本来提高系统整体的吞吐量。

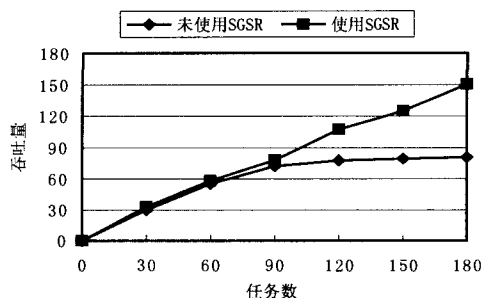


图 5 吞吐量比较

结束语 本文提出的 SGSR 网格服务复制机制,实现了有状态的网格服务的复制。网格副本服务技术只能进行无状态的副本复制,这是二者最大的区别。SGSR 完成了对网格服务所需硬件环境的搭建,软件环境的复制,网格服务程序代码、有状态资源的状态和内存状态的复制。SGSR 的注册和网格副本服务技术的注册机制大致一样,可以使用网格副本服务技术的注册机制进行注册。SGSR 在所需资源是安装态的情况下,能够在很短的时间内完成复制任务,为网格服务的复制提供了一种参考的方法。SGSR 能够很好地解决网格服务的过载和访问瓶颈问题,在达到某些特定条件时能够主动对网格服务进行复制。在以后的工作中,我们会继续完善安全的功能,并考虑当原网格服务进行更新时复制的新服务如何进行更新。

参考文献

- [1] Keahey K, Foster I, Freeman T, et al. Virtual Workspaces in the Grid//Europar. Lisbon, Portugal, 2005
- [2] Keahey K, Foster I, Freeman T, et al. Virtual Workspaces: Achieving Quality of Service and Quality of Life in the Grid. Scientific Programming Journal, 2006
- [3] Krsul I, Ganguly A, Zhang J, et al. VMPlants: Providing and Managing Virtual Machine Execution Environments for Grid Computing//SC04. Pittsburgh, PA, 2004
- [4] Foster I, Frey J, Graham S, et al. Modeling Stateful Resources with Web Services. <http://www-128.ibm.com/developerworks/webservices/library/specification/ws-resource/ws-modelingresources.html>
- [5] <http://www.globus.org>