获“西安石油大学优秀学术著作出版基金”资助

数据网格信任管理技术

DATA GRID TRUST MANAGEMENT TECHNOLOGY

张仙伟 著

中国石化出版社
HTTP://WWW.SINOPEC-PRESS.COM

获“西安石油大学优秀学术著作出版基金”资助

数据网格信任管理技术

张仙伟　著

中国石化出版社

内 容 提 要

本书以数据网格、信任管理技术为研究对象，介绍了数据网格环境下的信任管理体系框架(TMAMG)和基于域的数据网格信任管理模型(DBTM－MG)，并在此基础上构造了基于群组－活动的访问控制模型，设计了基于域的信任评估算法，最后基于数据网格访问控制模型，应用开源网格工具(Globus Toolkit)开发并实现了信任管理原型系统。

本书可供计算机科学与技术、软件工程、信息安全研究领域的科研人员、工程技术人员以及高等院校相关专业师生阅读和参考。

图书在版编目(CIP)数据

数据网格信任管理技术 /张仙伟著. —北京：中国石化出版社，2016.5
ISBN 978－7－5114－4064－8

Ⅰ.①数… Ⅱ.①张… Ⅲ.①网格—安全技术 Ⅳ.①TP393.08

中国版本图书馆 CIP 数据核字(2016)第 112096 号

未经本社书面授权，本书任何部分不得被复制、抄袭，或者以任何形式或任何方式传播。版权所有，侵权必究。

中国石化出版社出版发行
地址：北京市东城区安定门外大街 58 号
邮编：100011 电话：(010)84271850
读者服务部电话：(010)84289974
http://www.sinopec-press.com
E-mail：press@sinopec.com
北京富泰印刷有限责任公司印刷
全国各地新华书店经销
*
700×1000 毫米 16 开本 8.5 印张 153 千字
2016 年 6 月第 1 版 2016 年 6 月第 1 次印刷
定价：38.00 元

前　言

信任管理是描述和解释安全策略、获取安全凭证、判断安全凭证是否满足相关安全策略的技术方法，可用于信息安全事前预防。数据网格作为一个应用网格，具有分布性、异构性、共享性和协同性等特点，在安全管理方面比其他网格平台有着更高的要求。在数据网格资源的安全管理中使用信任管理技术，对于促进制造企业建立良好、可信的网络合作环境，提高资源共享、制造协同的安全水平具有重要意义。

本书围绕数据网格信任管理技术及其应用展开了深入探讨，主要研究内容与创新点如下：

(1)根据制造业信息化对数据网格信任管理技术的需求，提出了一个数据网格环境下的信任管理体系框架(TMAMG)，其组成为：信任基础知识定义层、信任存储管理层、信任初始化层、综合信任评估层、直接信任评价层、信任更新层、角色管理模块层、信任策略管理层和信任日志管理、应用接口层。该框架为信任管理研究提供了一个可行的参考标准框架，为开发数据网格中信任管理系统奠定了理论基础。

(2)依据数据网格环境下域结构的显著特征，结合 TMAMG 体系框架，提出了基于域的数据网格信任管理模型(DBTM－MG)。该模型由信任评估、信任评价和信任更新三大部分组成。它基于二层域结构管理机制，依据制造业的行业划分特点建立上层域，它不受地域的限制，方便实现动态扩展，并且能够满足不同行业、不同管理的要求；然后根据合作域普遍存在的特征建立下层域，正确地模拟出资源实体间信任建立的过程。该模型可以满足制造业信息化中跨多种行业合作的信任管理需求，并且能够反映出角色对信任管理的影响。

(3)构造了数据网格信任管理环境中基于群组－活动的访问控制模型。该模型基于层次 RBAC 模型进行扩展，用活动对角色和权限对象进行封装，并增加了活动状态、活动层次和活动依赖，从而实现动态的权限控制和灵活的权限粒度，符合数据网格中多企业协同的业务流程动态变化和操作依赖的要求。用群组和群组层次表示参与协同的组织中原有的组织结构和访问控制机制，从而保持了参与协作的组织原有结构的自治性及权限的分布式管理。此外，在协同过程中为群组提供类 Unix 的系统默认最小权限配置，从而简化了管理员的权限管理工作。

(4)为了实现直接信任的合理评价，采用构建主观层次型信任评价指标树技术，基于 DBTM－MG 信任管理模型，设计了基于域的信任评估算法。该算法采用了 D－S 理论中串联运算和并联运算的规则，利用求积方式实现单条路径的推荐信任求解，利用合作角色和行业角色作为因子的加权算法实现了多条路径的综合求解。该算法可有效避免推荐资源节点的单节点权值过重和受待遇不公平的现象，实现了信任值计算。

(5)基于以上理论研究成果，针对数据网格协同的需求，在数据网格访问控制模型的基础上，应用开源网格工具(Globus Toolkit)开发并实现了信任管理原型系统。该原型系统由基础参数选择、信任收集、信任评价、信任更新和数据网格访问控制五大子系统组成。

应用本书的研究成果，在数据网格资源的安全管理中使用信任管理技术，可提高制造业信息化资源共享和制造业网上协同的安全水平，促进制造业信息化的快速发展。

本书共分为七章。其中，第一章是数据网格的概述及其发展现状；第二章通过对数据网格信任管理技术的背景分析，根据信任基本组件和数据网格的特点，提出了数据网格信任管理框架；第三章是访问控制机制；第四章介绍了信任评估的研究现状，并提出了数据网格环境下信任评估的关键问题，同时提出了基于域的信任评估策略；第五章是信任管理原型系统设计与实现；第六章是信任管理原型系统运行界面及性能分析；第七章是结论与展望。

由于笔者水平有限，书中不妥之处在所难免，敬请各位专家、学者批评指正。

此为试读，需要完整PDF请访问：www.ertongbook.com

目　录

第一章 绪 论

第一节 数据网格概述

一、网格概念

网格，在生物学中是由支柱和细层组成的网格状骨骼结构。在信息学中，网格是一种用于集成或共享地理上分布的各种资源(包括计算机系统、存储系统、通信系统、文件、数据库、程序等)，使之成为有机的整体，共同完成各种所需任务的机制。

网格是一种基础设施，是在动态、多机构的虚拟组织中进行协同资源共享和问题解决的技术。通过网格，可把整个互联网上分散的、闲置的计算资源整合成一台统一的、开放的、巨大的虚拟超级计算机，从而帮助用户实现分布资源的全面共享和问题的协同求解；可对多个组织所拥有和管理的高性能计算机、网络、数据库以及科学工具进行综合、协同使用。

网格技术是在解决对等资源的共享和解决动态、分布式的虚拟组织所遇到的问题的过程中发展起来的。第一代 Internet(互联网)通过 TCP/IP 协议实现了计算机硬件的连通；第二代 Internet(互联网)通过 Web 技术实现了全球网页的连通；而网格试图实现互联网上所有资源的全面连通，包括计算资源、存储资源、通信资源、软件资源、信息资源和知识资源等。因此，网格技术又被喻为“第三代 Internet(互联网)”。

在全球化、信息化的时代背景下，制造企业在面对某个市场需求时，经常会围绕着新产品的开发，利用各自的管理体系、网络系统、软件系统等动态整合各自组织或企业间的优势资源，形成企业联盟，进行网上协同合作。但是，基于 Intranet/Internet 的网络协作平台没有采用开放的体系结构作支撑，没有通用的标准和规范作基础，限制了企业间各自系统的互操作，导致一些企业内部及企业之间存在着信息“孤岛”，只能采用紧耦合的方式实现资源的共享与互操作。然而，

网格技术为解决这些问题提供了有效的理论与技术支撑。

网格是一种新兴的技术，正处在不断的发展和变化当中。学术界和商业界围绕网格开展的研究有很多，其研究的内容和名称也不尽相同，因而网格尚未有精确的定义和内容定位。比如国外媒体常用“下一代互联网(NGI)”、“Internet2”、“下一代 Web”等来称呼网格相关技术。但“下一代互联网(NGI)”和“Internet2”又是美国的两个具体科研项目的名称，它们与网格研究目标相交叉，研究内容和重点有很大不同。企业界用的名称也很多，有内容分发(Contents Delivery)、服务分发(Service Delivery)、电子服务(E-Service)、实时企业计算(Real-Time Enterprise Computing，简称 RTEC)、分布式计算(Peer-to-Peer Computing，简称 P2P)、Web 服务(Web Services)等。中国科学院计算所所长李国杰院士认为，网格实际上是继传统互联网、Web 之后的第三次浪潮，可以称之为第三代互联网应用。

网格是利用互联网把地理上广泛分布的各种资源(包括计算资源、存储资源、带宽资源、软件资源、数据资源、信息资源和知识资源等)连成一个逻辑整体，就像一台超级计算机一样，为用户提供一体化信息和应用服务(计算、存储、访问等)，虚拟组织最终实现在这个虚拟环境下进行资源共享和协同工作，彻底消除资源“孤岛”，最充分地实现信息共享(图 1-1、图 1-2)。

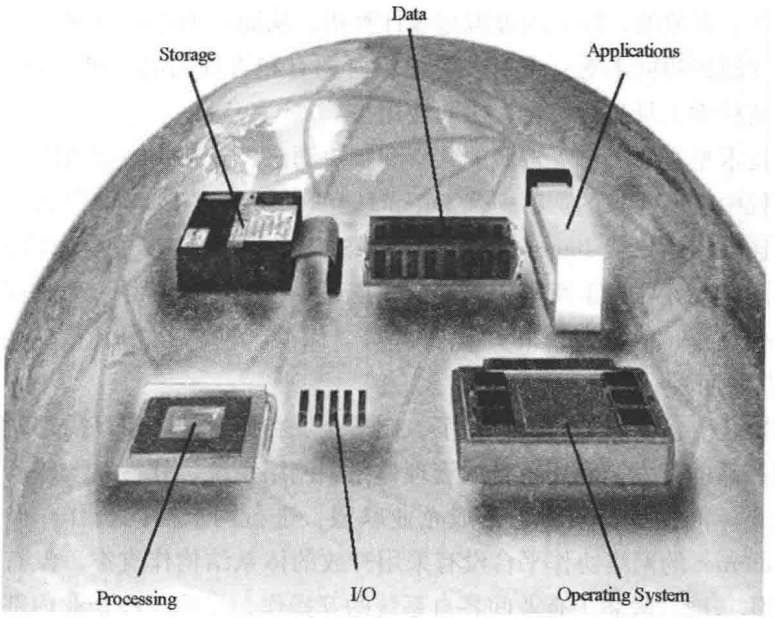

图 1-1　网格——把因特网变成一台虚拟计算机

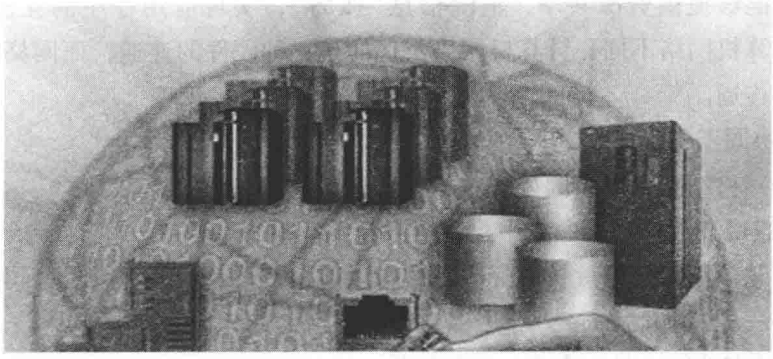

图 1-2 网格计算机

二、网格特点

在介绍网格的特征之前，我们首先要解决一个重要的问题：网格是不是分布式系统？这个问题之所以必须回答，因为人们常常会问到另一个相关的问题：为什么我们需要网格？现在已经有很多系统(比如海关报关系统、飞机订票系统)实现了资源共享与协同工作。这些系统与网格有什么区别？

对这个问题的简要回答是：网格是一种分布式系统，但网格又不同于传统的分布式系统。IBM Global Service 与 EDS 是在这个分布式领域最著名的公司。构建分布式系统有三种方法：传统方法(我们称之为 EDS 方法)、分布自律系统(Autonomous Decentralized Systems，简称 ADS)方法、网格(Grid)方法。分布自律系统(ADS)通常用于工业控制系统中。网格方法与传统方法的区别见表 1-1。

表 1-1 网格方法与传统方法的区别

特 征	传统分布式系统	网格
开放性	需求和技术有一定确定性、封闭性	开放技术、开放系统
通用性	专门领域、专有技术	通用技术
集中性	很可能是统一规划、集中控制	一般而言是自然进化、非集中控制
使用模式	常常是终端模式或 C/S 模式	服务模式为主
标准化	领域标准或行业标准	通用标准(+行业标准)
平台性	应用解决方案	平台或基础设施

通过以上对比，网格方法具有以下四点优势：

(1)资源共享，消除资源“孤岛”。

网格能够提供资源共享，消除信息“孤岛“，实现应用程序的互连、互通。网格与计算机网络不同，计算机网络实现的是一种硬件的连通，而网格能实现应用层面的连通。

(2)协同工作。

很多网格结点可以共同处理一个项目。

(3)通用开放标准，非集中控制，服务质量很高。

网格基于国际的开放技术标准，这与以前很多行业、部门或者公司推出的软件产品不一样。

(4)动态功能，具有高度的可扩展性。

网格可以提供动态的服务，能够适应变化。同时，网格并非是限制性的，它具有高度的可扩展性。

三、网格体系结构

网格体系结构是关于如何建造网格的技术，包括对网格基本组成部分和各部分功能的定义和描述、网格各部分相互关系与集成方法的规定、网格有效运行机制的刻画。显然，网格体系结构是网格的骨架和灵魂，也是网格最核心的技术，只有建立合理的网格体系结构，才能够设计和建造好网格，才能够使网格有效地发挥作用。

伊安·福斯特(Ian Foster)将网格体系结构定义为“划分系统基本组件，指定系统组件的目的与功能，说明组件之间如何相互作用的技术”。网格体系结构包括两个层次的作用：一是标识出系统的组成部分，清晰地描述出各部分的功能、目的和特点；二是描述各组成部分间的关系，以及如何将各部分有机地结合在一起，形成完整的网格系统，也就是将各个部分进行集成的方式或方法。

主流的三个网格体系结构为：

(1)伊安·福斯特(Ian Foster)等提出的五层沙漏结构(Five-Level Sandglass Architecture)。

(2)在以 IBM 为代表的工业界的影响下，考虑到 Web 技术的发展与影响，伊安·福斯特(Ian Foster)等结合五层沙漏结构和 Web 服务(Web Services)提出的 OGSA(Open Grid Services Architecture，开放网格服务体系结构)。

(3)由 Globus 联盟、IBM 和 HP 于 2004 年初共同提出的 WSRF(Web Service Resource Framework，称为 Web 服务资源框架)。WSRF v1.2 规范于 2006 年 4 月被批准为 OASIS (Organization for the Advancement of Structured Information

此为试读，需要完整PDF请访问：www.ertongbook.com

Standards，称为结构化信息标准促进组织)标准。

网络计算体系结构的发展可以概括为：

(1)层次体系结构。

(2)开放网格服务体系结构 OGSA。

(3)OGSI→WSRF(WS - Resource Framework)(图 1-3)。

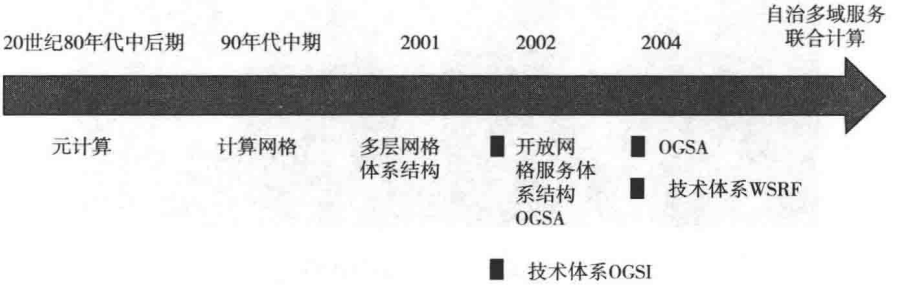

图 1-3 网络计算体系结构的发展

网格存在的主要问题为：在动态变化的、多机构组成的虚拟组织(VO)内的协作资源共享和问题求解。

(1)允许分布的服务和资源集成。

(2)采用通用的协议和基础支撑。

(3)获得较好的 QoS 服务(图 1-4)。

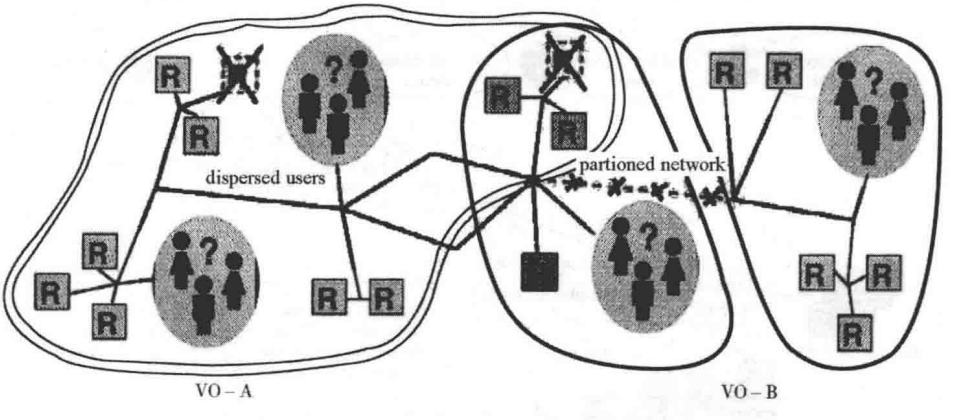

图 1-4 网格存在的主要问题

接下来，举一个虚拟组织的例子：

(1)CERN's Large Hadron Collider(图 1-5)。

(2)1800 Physicists, 150 Institutes, 32 Countries(图 1-6)。

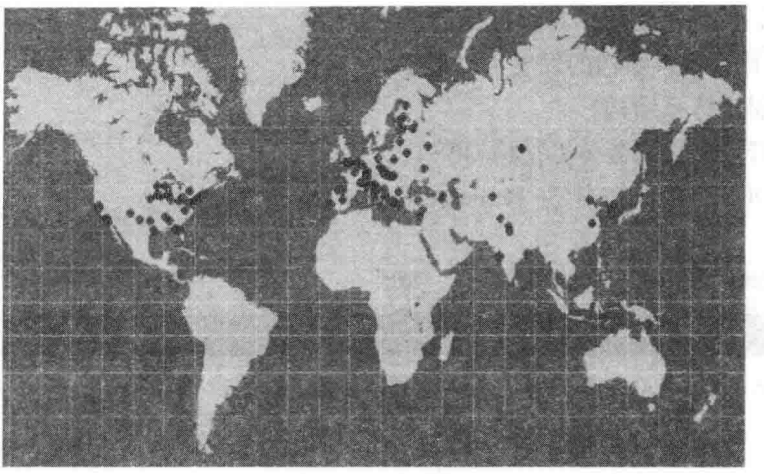

图 1-5 CERN's Large Hadron Collider

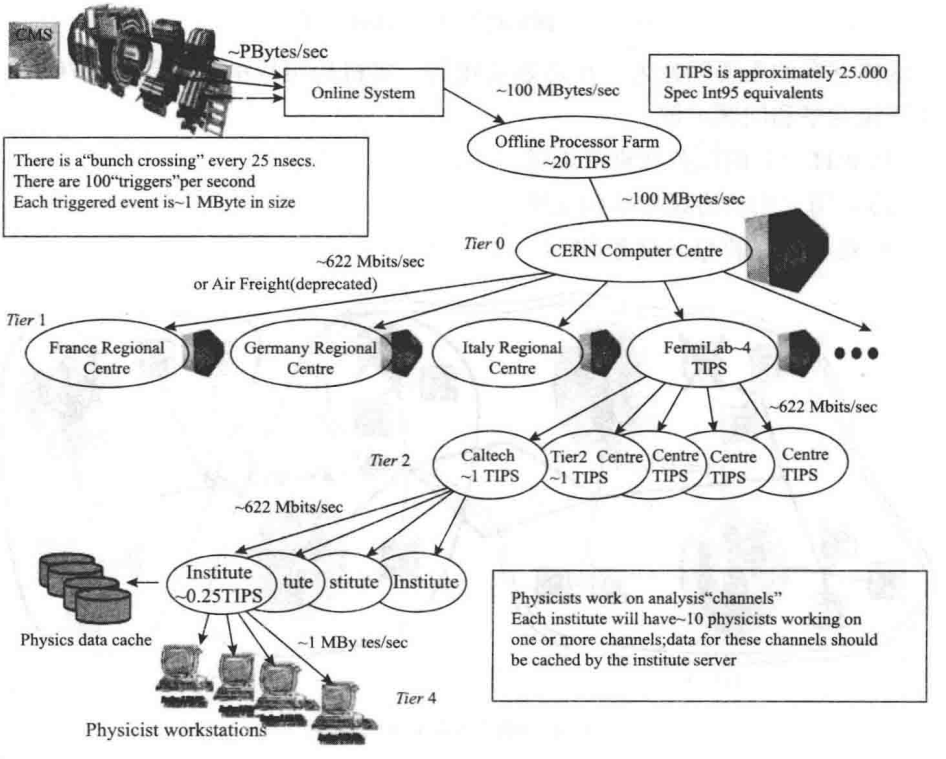

图 1-6 Grid Communities & Applications: Data Grids for High Energy Physics

五层沙漏体系结构是由 Ian Foster 提出的一种具有代表性的网格体系结构，其影响较大，特点就是简单，主要侧重于定性描述而不是具体的协议定义，容易

从整体上进行理解。五层沙漏体系结构最基本的思想就是：以协议为中心，强调服务与 API 和 SDK 的重要性。

五层沙漏结构的设计原则是要保持参与的开销最小，即作为基础的核心协议较少，类似于 OS 内核，以方便移植。

另外，沙漏结构管辖多种资源，允许局部控制，可用来构建高层的、特定领域的应用服务，支持广泛的适应性。

在五层结构中，资源层和连接层共同组成了瓶颈部分，使得该结构呈沙漏形状。

其内在含义就是各部分协议的数量是不同的。其最核心的部分能够实现上层各种协议向核心协议的映射，同时实现核心协议向下层各种协议的映射，核心协议在所有支持网格计算的地点都应该得到支持，因此核心协议的数量不应该太多，这样核心协议就形成了协议层次结构中的一个瓶颈(图1-7～图1-9)。

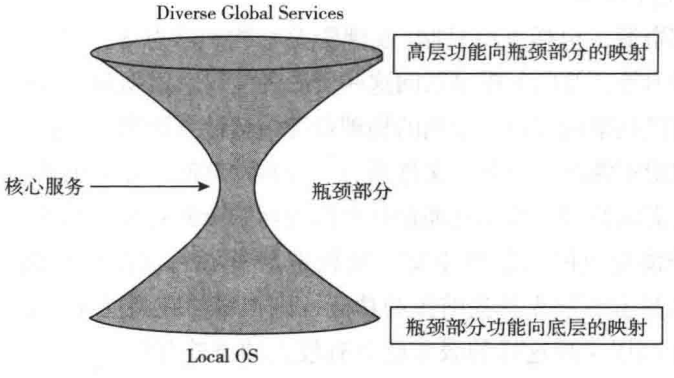

图 1-7 沙漏结构设计图

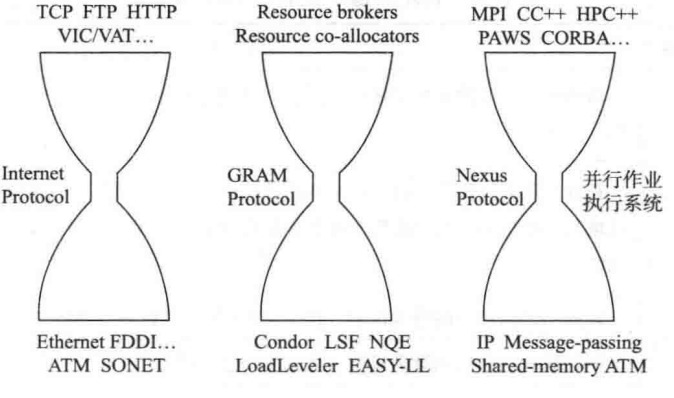

图 1-8 沙漏图

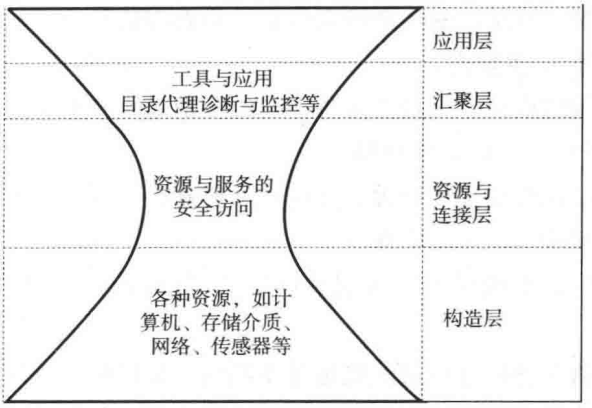

图 1-9　网格沙漏结构示意图

五层沙漏结构主要包括以下五个层次：

(1)构造层(Fabric)。

控制局部资源，包括查询机制(发现资源结构和状态等信息)、控制服务质量的资源管理能力等，并向上提供访问这些资源的接口。由物理或逻辑实体组成，目的是为上层提供共享的资源。常用的物理资源包括计算资源、存储系统、目录、网络资源等；逻辑资源包括分布式文件系统、分布计算池、计算机群等。构造层组件的功能受高层需求影响，基本功能包括资源查询和资源管理的 QoS 保证。

构造层资源提供的功能越丰富，则构造层资源可支持的高级共享操作就越多，例如，如果资源层支持提前预约功能，则很容易在高层实现资源的协同调度服务，否则在高层实现这样的服务就会有较大的额外开销。

特定构造层资源及其功能特性见表 1-2。

表 1-2　特定构造层资源及其功能特性

构造层资源举例	功能特性
计算资源	启动程序、监控和控制进程的执行、控制进程资源分配的管理机制、提前预留机制、查询功能
存储资源	存放与获取文件的机制、第三方高性能传输方式、读写文件子集机制、远程数据选取与归纳机制、对分配用于数据传输资源的控制管理机制、提前预约机制、查询功能
网络资源	对网络传输资源的管理机制、查询功能(用来得到网络特性和负载)
代码库	源代码和目标代码管理机制，比如 CVS 控制系统
目录	目录查询与更新操作机制，比如关系数据库

(2)连接层(Connectivity)。

支持便利安全的通信，该层定义了网格中安全通信与认证授权控制的核心协议，用于网格的网络事务处理。通信协议允许在构造层资源之间交换数据，要求包括传输、路由、命名等功能。实际上，这些协议大部分是从 TCP/IP 协议栈中抽取出来的。资源间的数据交换和授权认证、安全控制都在这一层控制实现。该层组件提供单点登录、代理委托、同本地安全策略的整合和基于用户的信任策略等功能(图 1-10)。

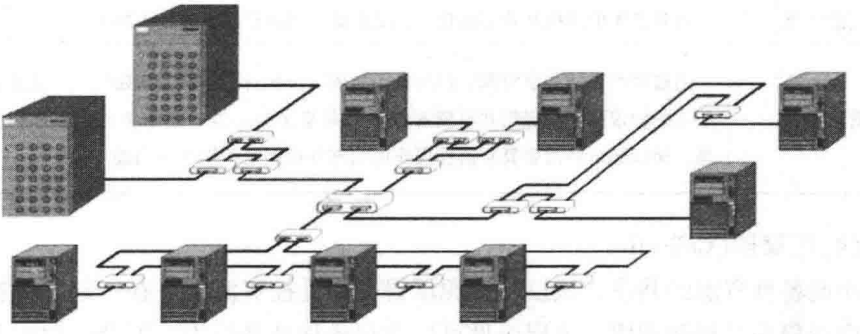

图 1-10 连接层功能

连接层安全认证特性见表 1-3。

表 1-3 连接层安全认证特性

特 点	描 述
单一登录	用户只需登录一次，就访问不同的构造层网格资源，不需要对不同的资源多次重复登录，也不需要用户进一步介入
代理	程序能访问用户认证的不同资源，还能够将它的部分权限授予另一个程序(受限制的代理)
与局部安全方法的集成	不同的资源可使用其局部的安全方案，但网格安全方案必须与那些局部的方案进行互操作，不要求网格安全方案完全代替局部安全方案，但它必须能实现向局部安全的映射
基于用户的信任机制	用户可使用多个提供者提供的资源，但并不要求资源提供者在安全环境中协同操作或互操作，即如果一个用户有权使用站点 A 和站点 B 的资源，用户能将站点 A 和站点 B 的资源结合起来使用，并不要求站点 A 和站点 B 的安全管理相互作用

(3)资源层(Resource)。

共享单一资源，该层建立在连接层的通信和认证协议之上，满足安全会话、

资源初始化、资源运行状况监测、资源使用状况统计等需求，通过调用构造层函数来访问和控制局部资源。资源层定义的协议包括安全初始化、监视、控制单个资源的共享操作、审计以及付费等。它忽略了全局状态和跨越分布资源集合的原子操作。定义了一些对单个资源的共享操作协议。

资源层的协议类型与描述见表1-4。

表1-4 资源层的协议类型与描述

协议类型	描 述
信息协议	得到资源的结构和状态信息，比如配置、当前负载、使用策略等
管理协议	通过判断访问共享资源，指出资源需求以及执行的操作，初始化共享关系，保证要求的协议操作与底层共享资源提供的共享策略一致，还要考虑记账和付费的问题，协议还可能需要具有监控操作的状态并控制某些操作的功能

(4)汇聚层(Collective)。

协调各种资源的共享，该层将资源层提交的受控资源汇集在一起，供虚拟组织的应用程序共享和调用。该层组件可以实现各种共享行为，汇聚层协议与服务描述的是资源的共性，包括目录服务、资源协同分配和调度以及代理服务、资源监测诊断服务、数据复制服务、网格支持下的编程系统、负荷控制、账户管理、负载管理系统与协同分配工作框架、软件发现服务、协作服务等功能。由此，说明了不同的资源集合之间是如何相互作用的，但不涉及资源的具体特征(图1-11)。

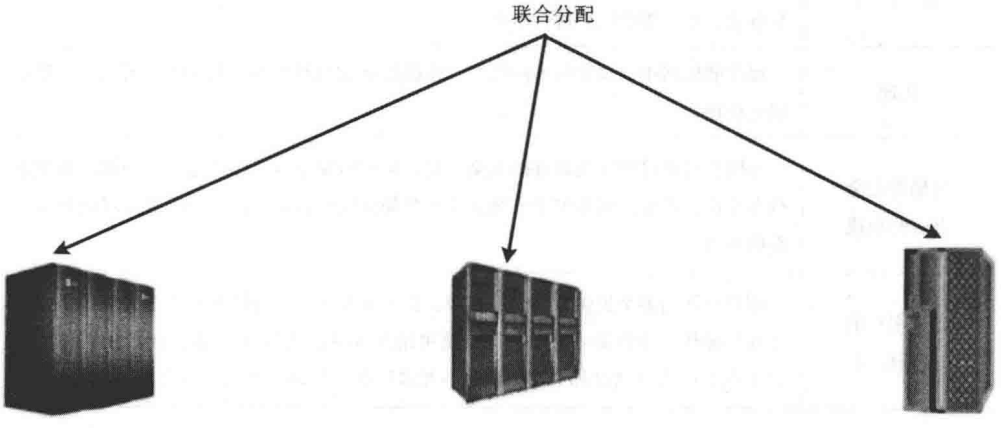

图1-11 汇聚层功能

此为试读，需要完整PDF请访问：www.ertongbook.com