

# 基于信息网络的中小型高校图书馆 数字资源整合研究

孙 丽

(黑龙江科技学院嵩山校区图书馆, 黑龙江 哈尔滨 150090)

**摘 要:** 中小型高校图书馆通过整合和利用现有的数字资源, 可以克服现有困难, 为教学科研提供高效优质的信息服务。通过网格技术, 以现有主流网格中间件 Globus Toolkit 为基础开发平台, 以资源整合和共享为目的, 建立基于信息网络的中小型高校图书馆数字资源整合平台的体系结构模型。对中小型高校图书馆的信息资源整合具有启发意义。

**关键词:** 中小图书馆; 数字资源; 资源整合; 信息网格; Globus Toolkit

中图分类号: G251

文献标识码: A

文章编号: 1002-1248 (2014) 01-0028-05

## Research on Digital Resources Integration of Small- and Medium-Size College Libraries Based on Information Grid

SUN Li

(Songshan Campus Library of Heilongjiang Institute of Science and Technology, Haerbin 150090, China)

**Abstract:** Some problems in digital resource construction of small and medium size college libraries are pointed out. It is suggested that integrate and utilize existing digital resources and provide highly effectiveness information service of teaching and scientific for the university users. Integration and sharing of resources for the purpose of grid technology becomes the mainstream of the information resource integrated technology, The paper firstly introduces the relevant technology of information grid, and Globus Toolkit 4, the existing grid network middleware is chosen as based exploit platform. the framework for the digital resources integration platform of small and medium size college libraries based on information grid is proposed.

**Keywords:** Small and Medium Size College Libraries; Digital Resources; Resource Integration; Information Grid; Globus Toolkit

### 1 概述

对于“中小型高校”的界定目前国内外尚未提出统一的标准。通常, 界定一所高校是不是中小型高校, 主要以在校生的规模为参考依据。一般说来, 在校生在 5 000 人以下的是小型高校, 5 000 人~15 000 万人是中型高校, 1.5 万人以上为大型高校。在市场经济条件下, 中小型高校是高等教育发展的主要力量<sup>[1]</sup>。教育部《2011 年全国教育事业统计公报》指出, 2011 年, 全国各类高等教育总规模达到 3 167 万人, 高等教育毛入学率达到 26.9%, 中国进入了高等教育大众化的阶段。《公报》显示, 普通高等本专科院校在招生人数、毕业生人数等方面在国内高等教育体系中占了超过

90%的比例, 是国内高等教育体系中的主要力量。同知名的大校、名校相比, 这些普通本专科院校多是中小型高校。这些中小型高校主要为地方经济做贡献, 培养出数以千万计的各类实用型人才, 他们是中国经济和社会发展的主力军。2010 年发布的《国家中长期教育改革和发展规划纲要 (2010—2020 年)》中, 将 2020 年的高教毛入学率预期定为 40%。这就需要中小型高校和大校、名校共同协调发展。

然而, 中小型高校与大型高校相比大多存在国家拨款少或几乎没有、经费短缺、学院大量扩招、办学条件不好、人才匮乏等方面的问题, 而且这类高校大多不重视图书馆的建设, 导致这类高校图书馆面临投

收稿日期: 2013-06-03

作者简介: 孙丽 (1979-), 女, 本科, 管理学学士学位, 黑龙江科技学院嵩山校区图书馆, 馆员。

入资金不足、馆藏不丰富、馆员素质不高、服务手段落后以及设备更新缓慢等方面问题。尽管如此,中小型高校图书馆承担的职责、任务与大型大学图书馆确是相同的,其读者对信息资源同样有较高使用需求,甚至有更高的专深度要求。中小型高校图书馆如何利用有限的资源来满足学生读者的阅读需求和教师的教学科研需要已经成为亟待解决的问题。

数字化时代的到来为中小型高校图书馆提高信息服务能力和效率带来了新的希望和机遇。随着计算机建设、网络技术、通讯技术以及存储技术的迅速发展,图书馆馆藏文献的类型发生了巨大的变化,数字资源在馆藏中所占比例日益增多,已成为高校图书馆深化信息服务的重要方式。数字资源的特点是信息存储量大,检索途径多、无时间、空间障碍,可以更好地满足大学教学、科研和阅读的需求。但由于数字资源本身数据结构、存储格式和发布方式不同,数字资源也呈现出海量、异构、分布、动态、自治等特性<sup>[2]</sup>,给用户检索和有效获取带来了诸多不便,也给图书馆信息资源共建共享的目标制造了障碍。要为用户提供高效、优质的文献信息资源服务,客观上有必要对学校现有的有限数字资源进行有效整合,屏蔽不同数据库资源之间的异构性,建立一个统一的检索平台,实现数字资源的一站式服务,以提高数字资源的使用效率。笔者试图本着高效、快捷、节约的原则,通过应用信息网格技术将中小型高校图书馆数字资源进行整合,并建立界面友好、功能完善的网格门户,为用户提供各种后台资源的无缝访问。

## 2 数字资源整合

### 2.1 数字资源整合

数字资源(Digital Resources)主要是指将文字、图像、语音、影像等内容,运用数字化技术手段和信息技术进行整合运用的产品,不仅包括数字化的图书、期刊、各类数据库,还包括其它任何以数字化形式存在的资源<sup>[3]</sup>。

数字资源整合(Digital Resource Integration),就是将各种物理上分布、异构和多样化的数字资源通过运用各种集成技术和方法将它们透明、无缝地连接为一体,为用户提供一站式服务平台,并能方便地集成各种数字资源服务,包括统一检索、资源链接、身份认证、个性化服务等,从而在各自独立的数据库资源基础上动态构建一个虚拟的数字图书馆,将各种资源以统一的方式呈现给读者,实现信息资源的有序组织、快速定位和有效揭示<sup>[4]</sup>。数字资源整合的目的是解决信息

孤岛问题。

利用现有技术手段,对跨平台、跨数据库、跨内容的数字资源进行有效整合,使之形成具有集成检索功能的数字资源系统,并将这种数字化服务同图书馆传统工作相融合,形成一个知识一体化的信息服务网络。在读者体验的角度,经过有效整合的数字资源系统使用更加友好,整合过的检索入口能够化繁为简,在检索中兼顾查全率和查准率,从而降低图书馆服务成本,有效提高资源利用率和图书馆服务水平。

### 2.2 中小型高校图书馆数字资源现状

中小型图书馆常用的主要数字资源包括:电子图书数据库、电子期刊数据库、视频音频数据库;自建数据库、学科导航数据库、馆藏书目数据库、随书光盘数据库、多媒体课件等。服务对象主要是学校的学生、教师和部分科研人员。其中,商业化数据库采购是数字资源建设的主要方式,自建特色数据库是相对有实力的中小图书馆的数字资源重点建设方向之一。从整体来看,多数中小型图书馆目前仍处在数字资源建设和积累阶段,对数字资源的整合程度不高。经过多年的积累和发展,众多中小型高校图书馆的数字资源无论从质量上还是数量上都有了质的飞跃;与此同时,如何对大量的数字资源进行有效整合,向读者提供便捷、高效的使用就成为迫在眉睫的问题。

由不同开发者建设的数据库在编码结构、表达方式、存储格式、检索接口等方面存在很大的差异性。这种差异性导致读者在使用不同数据库的过程中,需要分别对其进行登录、检索等操作,甚至需要在不同数据库间进行多次切换,对检索结果进行查重、比较才能获得满意的数字资源。这是对读者时间和精力浪费,不仅有违“节省读者的时间”的精神,也必然会影响图书馆数字资源的有效利用。人类阅读活动中的主要矛盾已经不再是有限的阅读材料同读者日益增长阅读需求的矛盾,而是读者有限的精力同信息高速膨胀之间的矛盾<sup>[5]</sup>。因此,数字资源的整合势在必行。

数字资源整合可以实现不同数字文献资源之间的沟通;最大限度地保持知识体系的完整性;使用户获得高质量、可信赖的信息资源。由于其统一的用户交互接口,可以使检索效率及资源利用率提高、响应速度快,具有信息代理和信息推送功能,具有分类元数据管理功能,可以提高知识获取效率,增强图书馆的科研能力和创新能力<sup>[6]</sup>。

目前,已经有一些图书馆管理系统研发企业关注到这一问题,在图书馆管理系统中植入了数字资源整合模块。但是这些管理系统的引进、使用成本相对较

高,难以在中小型高校图书馆中进行普及。而网络技术由于其开发难度和运行维护成本较低、灵活性强等特点,可以在不大量增加中小型高校图书馆服务成本的情况下对数字资源进行整合。

### 3 网络技术概述与信息网格

#### 3.1 网络技术 (Grid Technology)

网格(Grid)技术是近年来兴起的一门信息新技术,代表了继Internet技术和Web技术之后的第三次浪潮,网格的构想源于电力网(power grid),其基本思想就是想借用电力网的“打开开关就能用上电,而不需要关心它从哪里来”的概念,将网上的所有资源整合起来,向用户提供按需获取的服务。它为人们提供了建设信息系统的新视角、新理念和新技术。其目标是要将互联网上的所有资源包括计算机、数据、设备和服务等动态集成起来,形成一个有机整体,“在动态变化的多个虚拟组织间,共享资源和协同解决问题”<sup>[7]</sup>。

美国阿岗(Argonne)国家实验室的首席科学家、美国网格计算项目的领导人之一伊安·福斯特(Ian·Foster),曾在1998年主编过一本书《The Grid: Blueprint for a New Computing Infrastructure》里这样描述“网格是构筑在互联网上的一组新兴技术,它将高速互联网、高性能计算机、大型数据库、传感器、远程设备等融为一体,为科技人员和普通老百姓提供更多的资源、功能和交互性。互联网主要为人们提供电子邮件、网页浏览等通信功能,而网格功能更多更强,能让人们透明地使用计算、存储等其他资源”<sup>[8]</sup>。

伊安·福斯特在What is the Grid? A Three Point Checklist一文中对网格的定义进行了界定,认为可以从以下三个方面判断一个系统是否为网格系统:一是要协调而非集中控制资源。也就是说网格要整合各种资源,协调各种使用者;二是要使用标准、开放、通用的协议和界面;三是要得到非平凡的服务质量。

#### 3.2 网络的主要特点

(1) 应用成本低<sup>[9]</sup>:利用因特网中的各种已有软、硬件资源支持各种应用,不仅可以节省应用成本,而且可以带动各种各学科研究水平的迅速提高。

(2) 分布性<sup>[10]</sup>:分布性是网格的一个最主要的特点。网格上的各类资源(硬资源和软资源)通常类型复杂、规模较大、跨越地理范围较广。

(3) 共享性:网格的根本特征是资源共享而不是它的规模。尽管网格资源是分散的,但是它们却是可以充分共享的。

(4) 自相似性:网格具有自相似性特征。网格的

局部和整体之间存在着一定的相似性,局部往往在许多地方具有全局的某些特征,而全局的特征在局部也有一定的体现。网格的自相似性在网格的建造和研究过程中有重要的意义。

(5) 动态性:随着时间的推移,网格拥有的资源或功能可以增加或减少。

(6) 异构性:网格可以包含多种异构资源,包括跨越地理分布的多个管理域。

(7) 多级管理域:由于构成网格计算系统的超级计算机资源通常属于不同的机构或组织,并且使用不同的安全机制,因此既需要各个机构或组织对其拥有的资源有自主的管理能力,又需要他们共同参与解决多级管理域的问题,以实现资源共享和互操作。

网格与因特网相比具有高性能、一体化、知识生产、协同工作、资源共享等技术优点。网格在应用领域体现出来的优势主要是高性能、低成本。

#### 3.3 网络的体系结构

网格体系结构是网格的骨架和灵魂,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。

网格的体系结构随着网格应用需求的增长不断演化,WSRF(Web Services Resource Framework)是现阶段较新的标准,此外五层沙漏结构和OGSA(Open Grid Services Architecture)同样是被寄予较高期望的体系。两者的主要区别在于其体系核心分别是“协议”和“服务”。

OGSA可以看做是“五层沙漏结构”与Web Service技术结合后的一种进化。网格技术和Web Service技术沟通支撑起了OGSA。“事事皆服务”和“事事为服务”是OGSA的灵魂核心。OGSI(Open Grid Services Infrastructure)对用户同网格服务间的互动进行了定义,是OGSA的核心规范,使其具体化、规范化。OGSI对向网格发送处理请求时所使用的接口进行了限定,是网格计算标准体系中的核心标准,其作用相当于Web服务中的WSDL(Web Services Description Language)。开放网格服务基础结构(OGSI),完成了将网格软件基础结构标准化的工作,定义了在网络环境中各种服务间使用的接口和协议的标准,为OGSA软件组件的交互操作性提供最大的空间和可能。OGSA的体系架构如图1所示。

(1) 资源层——构成网格能力的资源包括物理资源和逻辑资源。网格的目的就是要管理好网格中的各种资源,为需要资源的用户提供透明一致的访问接口。物理资源包括服务器、存储器、网络等硬件设施。





图1 OGSA 的体系架构

物理资源之上是逻辑资源。它们通过虚拟化和聚合物理层的资源来提供额外的功能。逻辑资源在物理网格之上提供了文件系统、数据库、目录和工作流管理等抽象服务。

(2) Web 服务层——包括节 Web 服务以及定义网格服务的 OGSi 扩展。架构中的第二层是 Web 服务。这里有一条重要的原则：所有网格资源（逻辑的与物理的）都被建模为服务。OGSi 规范定义了网格服务并建立在标准 Web 服务技术之上。OGSi 利用诸如 XML 与 Web 服务描述语言 WSDL 这样的服务机制，为所有网格资源指定标准的接口、行为与交互。

(3) 基于 OGSA 架构的服务层。Web 服务层及其 OGSi 扩展为下一层提供了基础设施：基于架构的网格服务。

(4) 网格应用程序层。随着时间的推移，丰富的基于网格架构的服务不断被开发出来，使用一个或多个基于网格架构的服务新网格应用程序亦将出现。

WSRF 作为 OGSi 重构和发展的产物，它基本保留了 OGSi 中的所有功能，同时又改进了一些语法，并在其表示中采用了不同的技术，使 OGSA 更灵活，更贴近于实现。

网格的架构决定于网格体系结构的设计，但不管采用何种体系结构，网格都必须具备资源管理、信息管理、数据管理、服务质量保证、安全等基本功能模块<sup>[1]</sup>。

### 3.4 信息网格

信息网格 (Information Grid) 在计算网格的基础上，利用现有的网络基础设施、协议规范、Web 和数据库技术，为用户提供一体化的智能信息平台。信息网格的建设目标是创建一种架构在 OS (Operation System) 和 Web 之上的基于互联网的新一代信息平台 and 软件基础设施。以实现分布式、智能化和协作性的信息处理，使用户可以通过统一入口访问所有信息。信息网格是软件集成和网格技术的结合。

在互联网站点中，数据和信息资源的分布分散性强，通过信息网格，使资源能够被统一管理、集中使用。通过网格门户 (Grid Portal) 整个网络资源透明的呈现在用户面前。用户看到的是一个逻辑门户上的若干与自己相关的频道，而不再需要从千上万个网站中搜索自己想要的信息。信息网格追求的最终目标是能够做到服务点播 (Service On Demand) 和一步到位的服务 (One Click Is Enough)<sup>[12]</sup>。这也是网格和 Web 技术的最大区别。信息网格为中小型高校图书馆建立跨越 Web 的资源整合和集成应用提供了新的平台。

## 4 中小型高效图书馆数字资源信息网格的构建

### 4.1 中小型高校图书馆数字资源整合模型的体系结构

基于目前中小型高校图书馆数字资源数据库分布零散，内部存在异构性，难以满足统一检索服务的实际需求，将信息网格技术引入到中小型高校图书馆中，以现有主流网格中间件 Globus Toolkit 4 为开发平台，构建了一种基于信息网格的中小型高校图书馆数字资源整合统一服务平台，从而使异构、分布的图书馆数字资源按照网格框架规范组织起来，构成逻辑上统一的资源整体。网格中间件作为网格技术的核心，其主要任务在于利用分布于整个互联网的异构资源，包括计算集群，存储设备，科学仪器等等，通过构成一个同构的环境使得这些资源能够为分布于各地用户提供协同式的服务，以达到在整个广域网范围内的资源共享。基于信息网格的中小型高校图书馆数字资源整合模型的体系结构如图 2 所示，该模型自下向上共分为五层，分别为资源层、连接层、网格中间件层、服务管理层、网格门户层。

### 4.2 各层次的主要内容

从图 2 可以看出，这几个层次紧密联系。

(1) 资源层：信息网格底层是资源层，它是网格中分布式资源的集合，主要包括各种类型的裸资源，各种异构的关系型数据库，文件系统，XML 文档，以及 WebServices 网络服务及机构的计算资源、存储资源、日益丰富开放获取 (OA) 资源等。

(2) 连接层：提供 SOA 的基本环境，提供服务描述、服务发现及服务调用的基本标准和规范。

(3) 网格中间件层 GT4：网格中间件层是网格体系结构的核心，它屏蔽了各种网络资源的异构性，支持共享资源的统一管理、分布调度与安全控制。网格中间件层分析网格资源层的资源信息和当前网格服务提交的任务要求，把任务分配到资源层不同的节点，

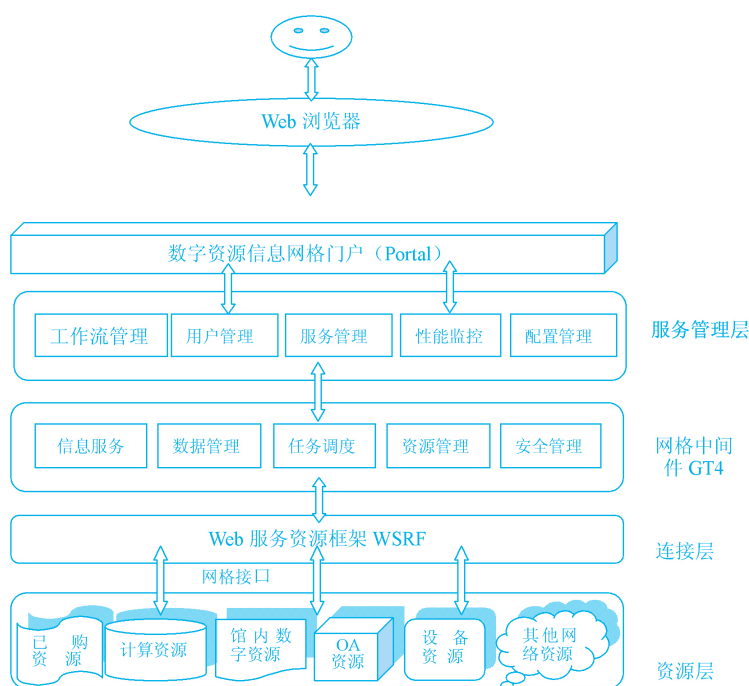


图 2 数字资源信息网格体系结构

由节点进行资源调度。在此选择的是比较著名且应用比较普遍的 Globus Toolkit4 作为网格中间件。

(4) 服务管理层: 该层实现了面向应用的高层服务, 提供了网格信息的基本功能; 而且还提供了对下一层服务的管理, 主要实现了服务的虚拟化功能以及资源信息的动态监控功能。另外, 它还可以为用户二次开发提供 API (Application Programming Interface) 支持。

(5) 网格门户层: 网格门户旨在方便用户与网格服务和网格资源的交互, 为网格服务和网格资源信息收集提供安全的环境, 并提供工具使用户能够应用网格服务和网格资源来完成自己的任务, 为用户提供一个网格软硬件资源视图和快速利用网格的计算能力; 保证用户能通过门户透明的访问网格的各项服务; 用户在网格门户层中设置要执行的任务, 网格门户层根据任务属性把任务发送到网格服务层, 并且最终把用户任务执行的结果显示给用户。

## 5 结语

高校图书馆作为社会信息服务机构的有机组成部分, 在区域经济社会高速发展的重要战略机遇面前应当大有可为<sup>[13]</sup>。时代在进步, 信息技术在飞速发展, 信

息网格的发展给信息资源管理带来的变化不仅仅是技术方面的, 而且对传统信息资源服务的概念和思路也有很大的影响, 网格技术的发展为中小型高校图书馆改变现状提供了机遇和技术支持。笔者对利用信息网格技术进行分布、异构数字资源整合进行了研究, 基于现有的网格中间件 Globus Toolkit 构建统一服务平台, 实现了数字资源的共享。该统一服务平台在功能上实现了网格环境下的数字资源整合, 用户不需要知道资源具体存放的位置, 就可以透明地访问多个网格数据资源, 为用户提供了一站式的检索服务。

## 参考文献:

- [1] 包和春. 必须重新认识中小型高校在高等教育发展中的战略地位—访北京联合大学商务学院院长孙明教授[J]. 北京教育: 高考版, 2005, (9): 14-16.
- [2] 孙雨生, 董慧. 国内信息网格技术研究进展(上)[J]. 现代图书情报技术, 200, (10): 27-31.
- [3] 刘雪娜, 陈朝. 我国数字资源产业化发展研究[J]. 情报科学, 2006, 24 (8): 1245-1248.
- [4] 王爱丽. 数字资源整合技术与模式研究[J]. 图书馆学研究, 2006, (3): 60-62, 73.
- [5] 蔡晨, 韩博文, 覃铭添. 高校图书馆利用读者资源开展阅读推广工作研究[J]. 情报探索, 2103, (3): 113-115, 119.
- [6] 王颖霞. 论数字图书馆的资源整合[J]. 现代图书情报技术, 2004, (S1): 106-107.
- [7] 朱敏. 基于网格的教育资源管理研究及模型设计[D]. 曲阜: 曲阜师范大学, 2006.
- [8] 张润彤, 樊宁. 网格就是商务[M]. 北京: 清华大学出版社, 北京交通大学出版社, 2006, (9).
- [9] 桂小林. 网格技术导论[M]. 北京: 北京邮电大学出版社, 2005, (1).
- [10] 应宏. 网格技术及其应用[J]. 计算机工程与设计, 2005, (10): 1685-1691.
- [11] 黄小龙, 潘大胜. 网格计算的优势及安全技术[J]. 网络安全, 2009, (2): 94-97.
- [12] 王明虎. 基于信息网格的决策支持系统研究[D]. 西安: 西安建筑科技大学, 2005.
- [13] 苏瑞竹, 张颖. 高校图书馆参与东盟信息服务机制探讨——以广西民族大学图书馆为例[J]. 农业图书情报学刊, 2013, (4): 153-156.