

校园网格文件系统的研究

马常霞

(淮海工学院 计算机科学系, 江苏 连云港 222005)

摘要:随着信息技术的发展,科学计算和并行技术所要处理的数据越来越多,现有的分布式文件系统已经越来越难以满足海量数据存储和资源在地理上的广泛协同和共享。校园网格文件系统是解决建立校园网格系统的重要技术之一,它将校园网格系统中的资源有机地统一起来,其应用将促进网格研究的进一步发展。

关键词:校园网格文件系统;目录服务;网格数据管理

中图分类号:TP393

文献标识码:A

文章编号:1005-3751(2006)04-0153-02

Research of Campus Grid File System

MA Chang-xia

(Department of Computer Science, Huaihai Institute of Technology, Lianyungang 222005, China)

Abstract: With the development of information technologies, data processed in scientific computation and concurrent technologies are more and more. Off-the-shelf distributed file system is difficult to satisfy mass data storage and resource cooperating with and sharing. Campus grid file system is one of the important technologies that integrates all resources in the campus grid system. Application of campus grid file system will speed the development of grid.

Key words: campus grid file system; directory service; grid data management

0 引言

早期的互联网应用带来了数据的大量集中,也为人们获取信息提供了迄今为止最为便捷的手段。然而各种信息的快速增长,使现有的 Web 信息服务器就像因特网中的一个孤岛。大量的信息被困在孤岛中,孤岛之间不能进行有意义的交流。因此人们希望更进一步地共享互联网上的一切资源,包括处理能力、存储容量、数据库、应用软件、文件等。解决上述问题的新技术正是“网格”。网格是因特网应用的新发展,又称为虚拟计算环境^[1]。网格把分布在各地的计算机连接起来,用户分享网上资源,如同使用一台超级计算机。简单地讲,传统因特网实现了计算机硬件的连通,Web 实现了网页的连通,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、知识资源等。

1 校园网格文件系统

1.1 网格文件系统

分布式文件系统的主要功能是支持用户在一定网络范围内,对一定数量的分布文件进行透明访问,典型的系统有: NFS(Network File System)^[2], AFS(Andrew File Sys-

tem)^[3]等。随着因特网的出现又有了面向因特网的分布式文件系统支持用户在因特网范围内对一定数量的分布式文件进行透明的访问。

网格文件系统主要建立在现有的数据管理系统的基础之上。它通常由两部分组成,一部分提供共享数据的定位和检索,利用元数据服务器对这些分布的、异构资源进行统一管理,使其形成一个整体,共同对外提供数据服务;另外一部分则提供数据共享协议,在异构系统之间进行文件、数据的共享。

1.2 校园网格文件系统的特点

校园网格文件系统作为网格环境中的一部分,它具有网格的共性,如分布性,文件和数据分布在校园网格系统内;动态性,文件和数据数量会随着网格规模的扩大而变化;多样性,校园网格文件系统中的资源种类繁多;虚拟性,网格文件系统中的文件和数据分布在不同的节点上,但是被组织在一起形成一个文件系统;规模不大、结构简单,校园网格系统中的资源相对于大型网格中的资源不仅数量少而且种类也少,这些资源多数以文件或计算机为主,资源种类少易于结构化。

1.3 校园网格文件系统的主要功能

(1)用户管理:主要完成收集用户信息,确认用户访问权限,这一部分主要由 LDAP(Lightweight Directory Access Protocol)^[4]服务器的安全机制来完成。

(2)资源管理:主要包括资源登记、删除、搜索等,它通常提供两种方式:目录行和图形界面。

收稿日期:2005-07-28

基金项目:江苏省教育厅自然科学基金资助项目(04KJD510019)

作者简介:马常霞(1975-),女,山西晋城人,讲师,硕士,主要研究领域为网络与分布式计算、移动 Agent 理论及应用。

(3)数据共享:包括资源定位,即用户对存储在目录信息服务中心的元数据进行搜索,检索到需要的数据;数据共享,在定位到资源后用户可以通过网格文件系统对资源进行共享。

(4)资源协同:大规模并行运算需要将任务分配到各个节点中,让它们协同工作,最终将结果发送到指定的地方。这就需要对参与计算的计算资源进行定位和计算数据的协同,包括中间结果的传递和最后结果的收集。

2 校园网格文件系统的总体框架

2.1 分析

文件系统一般包括目录信息操作和文件、数据操作两部分,与之相对应的网格文件系统包括网格目录信息服务,又称为元数据目录服务(Metacomputing Directory Services)和网格数据文件的共享操作,又称为网格数据管理两部分。目前广泛用于目录信息服务的是 LDAP 协议。网格数据管理主要包括分布式数据的访问和分布式计算资源的访问包括计算结果或计算数据的共享。

校园网格文件系统的构建主要基于 OGSA^[5](Open Grid Service Architecture)体系结构中面向“服务”思想,即网格就是可扩展的网格服务的集合,如图 1 所示,网格服务=接口/行为+服务数据。

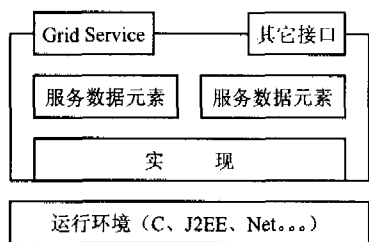


图 1 网格服务

整个校园网格文件系统功能实现由一组服务完成。这些服务大致可分解为 3 类:为用户提供文件系统的目录信息的元数据目录服务;目录访问模式服务实现,目录信息的获得与添加、删除、查询的操作;文件、数据的共享服务完成数据传输功能,传输的过程对用户透明。

2.2 总体框架

结合校园网格文件系统提供的 3 类服务,其总体框架如下:

(1)元数据目录服务提供虚拟的单一文件系统的映像。

(2)虚拟文件目录访问模式的实现由 4 部分组成:网格入口(Grid Portal)、客户端代理、服务器端代理、LDAP 服务器。

* 网格入口:提供给用户一个界面,令用户不仅可以浏览、操作本地文件系统,也可以对虚拟网格目录进行操作,如从服务器获取元数据目录信息;它也是用户访问服务器信息的中介,一方面它响应用户发出的请求并发送到服务器端代理,另一方面它将服务器端返回的信息转换为用户可以理解的格式;它还是数据共享的桥梁,当接到用

户数据共享的请求时,它会生成 FTP 客户端实例将请求发送到文件所在的对应客户端代理。

* 客户端代理:是客户端之间实现数据共享的桥梁。它在固定端口进行监听,响应别的客户端发送的数据共享请求,并生成 FTP 服务器端实例对其响应,形成数据传输链路。

* 服务器端代理:是客户端和服务端之间的桥梁。它响应客户端代理的用户请求,从本地提取数据,并传送给客户端代理。

* LDAP 服务器:它响应客户端代理的请求,在本地访问数据。

(3)实现数据和文件共享的模式为:用中心节点即 LDAP 服务器提供元数据目录服务;其余节点之间采用 P2P(Peer - Peer)模式,它是下一代互联网应用的主要框架形式之一,主要用于电子商务、分布式文件共享、分布式数据库方面。各节点之间完全对等,实现节点间的数据共享;节点间数据的传输采用因特网普遍使用的 FTP 协议。

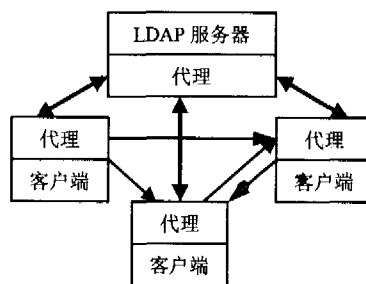


图 2 总体框架

在校园网格文件系统总体框架图中,双箭头线刻画了信息查询路径,它们总是在网格入口和服务器端代理之间进行;单箭头线刻画了数据的请求路径,它们总是在网格入口和另一端客户端代理之间进行。每一个客户端都可以看成是数据传送的发起者(请求方)也是响应者(服务方),所以说客户端之间是对等的。

3 结束语

网格计算环境可以连接广域范围内不同标准的异构“孤岛”,形成庞大的全球性计算体系,是因特网发展的高级形式。而作为校园网格系统中重要组件之一的网格文件系统的研究必将推进全球网格的进一步发展。随着对校园网格系统研究的不断深化,迫切需要完善校园网格文件系统,文中仅给出了校园网格文件系统构建的思路和框架,该系统的实现工作还有待于进一步展开。

参考文献:

- [1] Smarr L, Catlett C E. Metacomputing[J]. Communications of the ACM, 1992, 35(6): 45 - 52.
- [2] NFS: Network File System Protocol Specification RFC 1094 [EB/OL]. <http://www.ietf.com/rfc/rfc1094.txt>, 1989 - 03.

换后得到的状态集合,则查询处理树的模拟退火算法如下:

```

procedure CXSA()
  S = S0;
  T = T0;
  minS = S;
  while not(frozen) do
    while not(equilibrium) do
      S' = random state in neighbors( S );
      C = cost( S' ) - cost( S );
      If ( C ≤ 0 ) then S = S' ;
      If ( C > 0 ) then S = S' with probability e-C/T;
      If cost( S ) < cost(minS) then minS = S;
      T = reduce( T );
    return(minS);

```

4 算法的改进

当操作算子连接的数目不少于2时,代价函数的空间形状如图2所示。

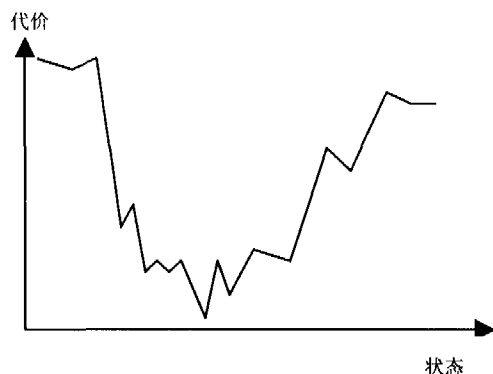


图2 代价函数的空间形状

从图2上看,代价函数空间形状属于一个“杯”形,在杯底区域包含了很多局部最小状态,需要多次局部最优优化才可以找到全局最小状态。在使用模拟退火算法进行优化时,能在杯底较快地找到全局的最小状态。但是温度 T 的初始值设置是影响模拟退火算法全局搜索性能的重要因素之一,如果初始代价选择在高代价区域,则搜索到全局最优解的可能性大,但是由于初始温度较高,要消耗一定的时间才可以稳定在杯底区域,因此要花费大量的计算时间,使得查询优化的效率较低。因此笔者对算法做出如下改进:

1)改进温度 T 的初始值设置方法。

在确定模拟退火算法的初始代价时,先经过一个重复的局部优化,找出一个较优的局部最小状态,并以此作为模拟退火算法的初值。

算法如下:

```

procedure jbcx()
  {minS = S0;
  while not (stopping_condition) do
    s = random state;
    while not (local_minim( s ))do
      S' = random state in neighbors( S );
      If cost( S' ) < cost( S ) then S = S' ;
    }
    if cost( S' ) < cost( S ) then minS = S' ;
    S0 = min( S );
  }

```

得到的 S_0 是局部最小优值,将其送至CXSA进行全局最优处理。这样可以加快优化的速度。

(2)改进温度管理问题。

考虑计算复杂度的切实可行性等问题,采用如下所示的降温方式:

$$T(t+1) = k \times T(t)$$

式中 k 为正的略小于1.00的常数, t 为降温的次数。

5 小结

在分布式环境下,由于数据库的分布性和自治性给数据的查询优化带来了很多不便,传统的查询优化方法不适合分布式环境下的查询优化。文中分析了分布式环境下全局查询优化的过程,并使用模拟退火算法去解决该过程。模拟退火算法具有质量高、初始鲁棒性强、通用易实现的优点^[2],但需要较长的计算时间,因此文中还提出了一种改进算法,通过改进模拟退火算法的初始值设置方法以及改进降温方法提高其计算的速度。

参考文献:

- [1] Silberschatz A, Korth H F, Sudarshan S. 数据库系统概念[M]. 杨冬青,唐世渭,等译. 北京:机械工业出版社,2002.
- [2] 熊平华. 基于网格计算平台的智能优化算法应用与研究[D]. 杭州:浙江大学,2004.
- [3] Steinbrunn M, Moerkotte G, Kemper A. Heuristic and Randomized Optimization for the Join Ordering Problem[J]. The VLDB Journal, 1997, 6(3): 8-17.
- [4] Selinger P G, Astrahan M M, Chamberlin D D, et al. Access Path Selection in a Relational Database Management System [A]. In Processing of the ACM SIGMOD Conf. on Management of Data[C]. Boston, USA: [s.n.], 1979. 8-17.
- [5] 王能斌. 数据库系统原理[M]. 北京:电子工业出版社, 2000. 116-135.

(上接第154页)

- [3] Carnegie Mellon University and IBM. AFS File System[EB/OL]. <http://www.cs.wisc.edu/csl/doc/howto/afs>, 1995.
- [4] Wahl M. Lightweight Directory Access Protocol (v3) [S]. RFC2251. 1997-12.

- [5] OGSA 结构描述, The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration (Draft 2.9, 6/22/2002)[EB/OL]. <http://www.gridforum.org/ogsi-wg/drafts/ogsa-drafts2.9-2002-06-22.pdf>. 2002-06.