

教育资源网格模型及副本创建策略*

邢长明⁺, 刘方爱, 杨林, 厉文

(山东师范大学 信息科学与工程学院, 山东 济南 250014)

Education Resource Grid Model and Replica Creation Strategies

XING Chang-Ming⁺, LIU Fang-Ai, YANG Lin, LI Wen

(School of Information Science and Technology, Shandong Normal University, Ji'nan 250014, China)

+ Corresponding author: E-mail: xingchm@tom.com, http://www.sdn.edu.cn

Xing CM, Liu FA, Yang L, Li W. Education resource grid model and replica creation strategies. Journal of Software, 2009,20(10):2844-2856. <http://www.jos.org.cn/1000-9825/3377.htm>

Abstract: The education resource grid is an effective measure to solve the problem of sharing distributed education resources at present. First, with regard to the education resource sharing in secondary and elementary schools, a hierarchy model of education resource grid is proposed and the functions of nodes in each layer are defined in this paper. By comparing with the European Data Grid, the education resource grid is analyzed with its characteristics. Then based on the hierarchy grid model, some factors, which influence the performance of replica creation strategy, are analyzed and a dynamic replica creation strategy (EDRS) is proposed by introducing two parameters, i.e., the network bandwidth and file size. With the data grid simulation tool OptorSim, the performances of EDRS with Caching-LRU, Caching-LFU and the economic model strategy are compared respectively. At last, the influence of different strategies on the performance of the education resource grid system is analyzed by synthesizing each parameter. The result indicates that the EDRS has a better system performance in applications of the education resource grid.

Key words: education resource; grid; replica creation; hierarchy; resource sharing

摘要: 教育资源网格是解决目前分布式教育资源共享问题的有效手段.针对中小学教育资源共享问题,提出了层次式的教育资源网格模型,定义了各层节点的功能.通过与欧洲数据网格对比,分析了教育资源网格的特点.基于层次式的教育资源网格,对影响副本创建策略性能的因素进行了分析,然后引入网络带宽和文件大小两个参数,提出了一种动态副本创建策略(dynamic replica creation strategy,简称EDRS).利用数据网格模拟工具OptorSim构建了教育资源网格虚拟环境,分析比较了EDRS策略与Caching-LRU策略、Caching-LFU策略和基于经济模型的副本创建策略的性能.最后,综合各项指标分析了不同策略对教育资源网格系统性能的影响.结果表明,EDRS策略在教育资源网格应用中有着更好的系统性能.

关键词: 教育资源; 网格; 副本创建; 层次式; 资源共享

中图法分类号: TP393 **文献标识码:** A

* Supported by the National Natural Science Foundation of China under Grant Nos.60373063, 90612003 (国家自然科学基金); the Shandong Provincial Natural Science Foundation of China under Grant No.Y2007G11 (山东省自然科学基金)

Received 2008-01-31; Revised 2008-04-02; Accepted 2008-05-05

网格作为新一代互联网的主要应用环境,为解决分布式资源共享问题提供了有效手段^[1].网格技术在 20 世纪 90 年代中期提出以后,研究重点集中在高性能计算网格、生物网格、交通网格和高校资源网格等方面^[1-4],然而,针对中小学教育资源共享的网格技术研究较少.目前,各地中小学都建立了自己的校园网络,并开发了大量的各种各样的教育资源,例如各学科的教学课件、实验案例、教学视频和教学管理软件等.这些资源的分布范围较广,应用面十分普及.但是目前,这些资源形成了一个信息孤岛,不同的学校很难共享,造成了资源的极大浪费.为了使现有资源得以充分利用,避免盲目地重复开发,如何利用网格技术共享这些资源是迫切需要解决的问题.

在资源分布式存储的前提下,如何组织、管理这些资源,实现教育资源的可用性,提高教育资源网格的生存性是我们关注的焦点.合理的资源组织模型是实现资源共享的前提.目前,针对各种不同的网络应用,提出了多种资源组织模型,其中具有代表性的有集中模型(如 Napster)、结构化模型(如 PAST)和非结构化模型(如 Gnutella).然而,这些模型并不适合于在行政区域和资源所有权方面有其独特性的中小学教育资源.服务质量是影响系统可使用性和生存性的关键.在网格系统中,创建副本是提高系统的服务质量的有效手段^[4].但是目前,针对副本创建策略的研究主要集中在以欧洲数据网格为代表的网格系统环境中^[4,5].教育资源网格与欧洲数据网格相比,在原始数据分布、网络带宽和节点存储能力等网络环境方面都存在着差异,这使得现有的副本创建策略不能应用于教育资源网格.

针对上述问题,我们主要做了如下工作:提出层次式教育资源网格模型,并对各层节点的具体功能进行了定义;针对教育资源网格的特点,分析了影响副本创建策略性能的关键因素;基于层次式教育资源网格,引入网络带宽和文件大小两个参数,提出了一种动态副本创建策略(dynamic replica creation strategy,简称 EDRS);利用数据网格模拟工具 OptorSim 构建教育资源网格虚拟环境进行仿真实验,比较了 EDRS 策略与 Caching-LRU 策略、Caching-LFU 策略和基于经济模型的副本创建策略的性能,综合分析了不同策略对教育资源网格系统性能的影响.结果表明,在层次式教育资源网格中,EDRS 策略能够降低用户延迟,减少副本创建次数,具有更好的系统性能.

1 相关研究

目前存在的分布式资源存储模型主要有结构化模型和非结构化模型.结构化资源共享系统(如 PAST)基于分布式哈希函数,资源严格地映射到节点,这类系统的查询路由比较高效,但却需要改变现有资源的存储模式,不能实现资源的自治性,而且这类系统存在拓扑结构维护代价较高和不支持模糊检索等局限性.非结构化资源共享系统(如 Gnutella)并不将数据存储到某个特定的节点,网络中的节点加入或离开所要遵循的规则也比较松散,然而其通过洪泛(flooding)机制完成消息的传递,这样就会产生大量的垃圾通信,占用大量的网络带宽,降低了互联网络的效率.考虑到中小学教育资源分散性的特点,按照资源存储在本地的原则,结合教育部门的组织模式,本文提出了层次式教育资源网格模型.

创建副本是提高网格系统的可扩展性、容错性、可用性和降低用户延迟的有效手段.就目前的研究来看,副本创建策略主要包括静态策略和动态策略.在数据量很大的网格环境中,由于用户的分散性和动态性,静态策略并不可行.网格系统需要动态策略,因为动态策略能够适应网格环境的动态变化^[4].目前存在的副本动态创建策略主要有:

文献[4]提出了 6 种不同的副本创建策略:No Replication, Best Client, Cascading, Plain Caching, Caching+Cascading, Fast Spread.文献[6]提出了两种副本创建策略:SBU 和 ABU.模拟实验结果表明,上述策略在大多数情况下都能降低用户访问延迟并节省网络带宽.然而, Cascading, Fast Spread, SBU 和 ABU 策略只适用于原始数据存储于顶层节点的层次式数据网格,并不适用于原始数据分散在底层的教育资源网格.其他几种策略虽然可适用于普通网络环境的教育资源网格,但也都只是简单地将用户请求文件的次数或频率作为创建资源副本的参数,它们并没有考虑到网络带宽和文件大小对用户延迟的影响.然而在实际网格环境中,请求次数最多的文件带来的用户延迟并不一定最大.

文献[5]提出了基于经济模型的副本创建策略,按照反向拍卖协议确定副本创建位置并进行副本选择,它将文件传输时间作为拍卖的价格指标.实验分析表明,该策略在欧洲数据网格环境中较好的性能.但是,由于该策略针对欧洲数据网格设计,并没有考虑各节点的存储能力,在节点存储能力较小时容易引起副本的频繁建立与删除.文献[7]提出了一种基于存储联盟的副本创建策略,考虑了单个节点的存储能力和用户访问文件的次数,然而它却忽视了网络带宽和文件大小对用户延迟的影响.除此之外,文献[8,9]针对不同结构的网格从不同角度提出了一系列动态副本创建策略,但是它们同样不适用于普通网络环境下的层次式教育资源网格.

上述分析表明,现有的副本创建策略要么针对特殊的网格环境设计,不能适用于普通网络环境的网格系统;要么考虑的网络参数较少,不能很好地反映实际的网格环境,造成副本的频繁建立与删除.另外,部分策略的实现也受到限制.本文针对一般网络环境下的中小学教育资源共享问题,首先提出层次式的教育资源网格模型,并对各节点的功能进行定义,然后基于该模型提出一种动态副本创建策略.该策略在副本创建时不仅充分考虑了网络带宽、文件大小和请求频率等因素,而且通过自底向上的副本创建方式解决了由于网格节点存储能力较小而容易引起副本频繁建立与删除的问题.

2 教育资源网格模型

教育资源网格其实是一种数据网格,它以实现教育资源的分布式共享为目标.现有的分布式资源共享模型要么需要严格的资源控制机制,不能实现资源的自治性;要么通过洪泛(flooding)机制进行消息传递,容易造成网络拥塞.为了改变这种情况,考虑资源特点、通信效率、组织模式等多种因素,我们提出层次式的教育资源网格模型.

2.1 层次式教育资源网格模型

考虑国家、省、市、县(区)的现行结构,教育资源网格采用层次式的结构模型,如图1所示.

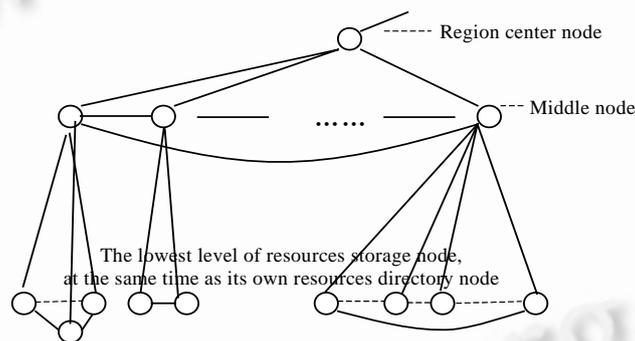


Fig.1 Hierarchical model of the education resource grid

图1 教育资源网格层次模型

该模型自顶向下形成层次式的树形结构.树的根节点为区域中心节点;中间节点为省、市、县的教育管理部门,可以根据系统规模进行扩展,为了便于描述,图1的中间节点仅有一层市级教育管理部门;树的叶节点为资源节点,它们代表各个学校,位于网格的最底层,负责本校资源的维护工作.各层节点的具体功能定义如下:

(1) 区域中心节点

中心节点维护整个系统的目录信息,以使用户对不同子域内资源的查询与访问;中心节点相当于一个中央服务器,负责全局管理,它通过分析各个地域资源的访问量找出热点资源,为用户延迟较高的热点资源建立副本;中心节点还负责响应中间节点的请求、本系统内的成员认证、计费管理等相关问题.

(2) 中间节点

中间节点维护其辖域内的分目录信息(关于目录信息结构本文不作探讨);中间节点根据副本创建策略存储

资源副本,并建立副本目录信息;中间节点响应低层节点的请求,资源节点请求资源时,中间节点首先查询本地资源目录信息,然后查询副本资源目录信息,当均不存在该资源时,向其上层节点发送请求。

(3) 资源节点

资源节点管理本地资源,包括资源的添加、删除和修改,实时地向上层节点报送本地资源的动态变化;存储和维护本地资源的目录信息;依据副本创建策略,存储资源副本,并建立本地副本目录信息;响应本地用户对资源的查询请求,当用户请求资源时,首先查询本地资源目录信息,再查询副本目录信息,当均不存在时,向其上层中间节点发起请求。

2.2 层次模型的优势

在现存的网格系统中,欧洲数据网格是最有名的采用层次模型的网格系统.教育资源网格为什么采用层次模型呢?首先,我国教育部门现行的组织模式(国家、省、市、县)自然构成一种层次式结构,因此采用层次式模型,使节点依据其所处的层次来管理其辖域内的资源,符合教育部门的管理模式;其次,采用层次模型,能够保证资源分布式存储在各自的学校,学校作为资源网格的底层节点,负责在本地维护本校的各种资源,能够实现资源的自治性,便于保护资源的知识产权;再次,根据文献[10]的分析,层次模型能够避免资源发现、资源检索以及资源传递过程中由于消息洪泛(flooding)引起的网络拥塞,提高通信效率;最后,在该层次模型中,兄弟节点之间建立环形结构,将环与层次结构结合,不仅具有层次结构的可扩展性,而且具有平面结构数据可用性和可靠性高的优点.另外,教育资源网格的层次模型并不同于欧洲数据网格的层次模型,下面对它们进行比较,以说明二者的区别。

2.3 与欧洲数据网格的比较

教育资源网格是一种数据网格,但它不同于欧洲数据网格.教育资源网格与欧洲数据网格相比,主要存在以下不同:欧洲数据网格以具有固定高带宽的网络环境为基础,而教育资源网格基于 Internet 和现有校园网,其网络带宽分布不均匀;欧洲数据网格的单个节点具有较强的计算能力和存储能力,而教育资源网格节点基于既有的 PC 机,单个节点的计算能力和存储能力都远远小于欧洲数据网格,并且节点所在的层次越低,其存储能力越差(例如,在计算机的配置上,市级节点一般低于省级节点);教育资源网格各个节点的响应速度不同(CPU 的主频、硬盘的读写速度),一般地,层次越高的节点其响应速度越快;特别是,欧洲数据网格的原始资源集中存储在顶层节点,而教育资源网格的原始资源分布式存储在底层节点。

以上网络环境的不同,使得欧洲数据网格的副本创建策略不能应用于一般网络环境下的中小学教育资源网格.特别是在欧洲数据网格副本创建策略中没有考虑节点的存储能力,如果在教育资源网格中采用其副本创建策略,则在节点存储能力较小时容易引起副本的频繁建立与删除.因此,如何寻找一种适合于教育资源网格模型的副本创建策略是我们讨论的重点。

2.4 副本创建策略的影响因素分析

教育资源网格用户基本操作主要包括数据文件的上传、检索和下载.在教育资源网格中引入副本机制后,合理地创建副本可以有效地提高用户资源检索和下载的速度,缩短用户请求资源的响应时间.下面,我们对影响教育资源网格副本创建的因素进行分析。

在教育资源网格中,一个数据文件可以有多个副本,每个副本具有一个物理文件名(PFN)和一个逻辑文件名(LFN).每个副本对应唯一的物理文件名,相同副本的逻辑文件名相同,即一个 LFN 对应多个 PFN.我们将一个 LFN 对应的所有 PFN 的集合记为|LFN|。

当用户对一个数据文件发出访问请求时,副本机制从该数据文件的多个副本中为用户选择一个响应时间最短的副本.因此,副本对特定请求的响应时间直接影响副本机制的性能.副本 PFN 对特定请求 Req 的响应时间 $ResponseTime(PFN, Req)$,可用下面的公式来计算:

$$ResponseTime(PFN, Req) = filetransfer_time(PFN, Req) + fileaccess_time(PFN, Req),$$

其中, $filetransfer_time(PFN, Req)$ 表示将文件 PFN 从其存储节点传输到请求节点的时间; $fileaccess_time(PFN,$

Req)表示访问文件 PFN 所花费的时间,包括传输前的等待时间、读写文件的时间等.

$filetransfer_time(PFN,Req)=PFN.size/transfer_rate$.其中, $PFN.size$ 表示副本的大小, $transfer_rate$ 为副本存储节点到请求节点间传输链路的最小带宽.因此,副本存储节点和请求节点之间的网络带宽以及副本大小直接影响到副本的响应时间.而影响 $fileaccess_time(PFN,Req)$ 的因素包括 PFN 存储节点的负载、存储节点的读效率和请求节点的写效率.一般认为, $filetransfer_time(PFN,Req)\gg fileaccess_time(PFN,Req)$ ^[7].

另外,用户的请求模式也直接影响着副本机制的性能.用户对数据的请求可能具有多种访问局部性,包括时间局部性和空间局部性等.一般地,用户访问数据的局部性越强,副本性能的表现就越明显.影响副本性能的因素还包括网络系统的拓扑结构以及单个节点的存储能力.教育资源网用户请求数据时通常先将数据从源节点复制到本地节点,然后在本地进行数据访问,当单个节点的存储能力较小时,容易引起副本的频繁建立和删除.

由以上分析可知,在设计副本创建策略时应注意以下几个因素:节点间的网络带宽、文件大小、用户的访问模式和节点的存储能力.下面基于这些因素探讨教育资源网的副本创建策略.

3 教育资源网动态副本创建策略

首先以一个具体例子进一步说明现有策略应用于教育资源网的问题.根据教育资源网的特点,在如图2所示的网格节点拓扑图中,假设节点 S_0 的存储能力为 100G,节点 S_1, S_4 的存储能力为 10G,节点 S_3, S_5, S_6 的存储能力为 1G;大小均为 500M 的数据文件 $lfn1, lfn2, lfn3$ 分别存储在 S_3, S_5, S_6 ;各节点间的网络带宽(Mbps)如图所示.假如用户向节点 S_2 分别以概率 0.6 和 0.4 提交作业 $job1$ 和 $job2$ 共 100 次, $job1$ 的执行需要文件 $lfn1, lfn2$, $job2$

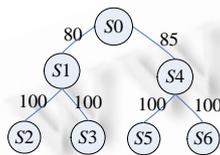


Fig.2 Site topological graph

图2 节点拓扑图

的执行需要文件 $lfn1, lfn3$.针对该例,在 S_2 的存储能力分别为 1G 和 600M 时,通过 OptorSim 模拟各种策略,其性能比较见表 1.

显然,随着 S_2 节点存储能力的减弱,Caching 策略和经济模型策略的性能都在降低.在节点存储容量为 600M 时,采用 Caching 策略,当 $job1$ 执行时将在节点 S_2 上创建 $lfn1, lfn2$ 的副本;当 $job2$ 执行时,将按照 LRU 或 LFU 策略进行空间置换,然后创建 $lfn1, lfn3$ 的副本.采用 LRU,不能保证创建时间最长的副本就是用户请求频率最少的;采用 LFU,不能保证用户请求频率最小的副本就是造成用户延迟最大的.因此, $job1, job2$ 的交替执行导致副本的频繁创建与删除.此时,如果将 $lfn3$ 的副本创建在节点 S_1 上,显然比在 S_2 上进行频繁的副本置换要好.采用文献[5]基于经济模型的创建策略,当 S_2 空间不足时,是否进行副本置换由其经济模型决定,显然比 LRU 和 LFU 要好.但由于没有考虑在节点 S_2 的邻居节点建立副本,当空间不足时,部分文件只能通过远程访问的形式获取,这显然也影响了系统性能.

Table 1 Performance index compare of different strategy

表 1 各种策略的性能指标比较

Strategy	Job execution time		Replica creation times		Network usage	
	1G	600M	1G	600M	1G	600M
No replica	4 454	4 454	0	0	1.0	1.0
Caching-LRU	1 083	4 454	49	200	0.245	1.0
Caching-LFU	1 116	4 454	46	200	0.23	1.0
Economic	1 066	3 096	3	4	0.205	0.72

如何减少副本创建次数和文件远程访问次数?我们主要通过两种手段解决该问题:一是在副本创建策略中引入副本代价和收益两项指标,只有收益大于代价时才进行副本置换;二是当低层节点不能创建副本时,判断其邻居节点是否可以创建副本.本文的 EDRS 策略包括 MEDRS(the main part of the dynamic replica creation strategy in education resource grid)策略和 AEDRS(the auxiliary part of the dynamic replica creation strategy in education resource grid)策略两部分,下面分别加以介绍.

3.1 动态副本创建主策略——MEDRS

教育资源网格由若干节点组成,节点间形成层次结构,节点记为 *site*。一个 *site* 可包含多个 *CE* 和 *SE*,其中:*CE* 为计算单元,用于处理用户的作业请求;*SE* 为存储单元,用于存储系统的原始资源和副本资源。每个节点维护一个历史访问记录,历史记录的一项为一个三元组,用(*SEID*,*PFN*,*frequency*)表示 *SE* 中的副本 *PFN* 被访问的频率,用(*CEID*,*LFN*,*frequency*)表示 *CE* 请求数据文件 *LFN* 的频率。

为了便于策略描述,我们引入以下定义和记号:

定义 1. 当用户对数据文件发出访问请求时,副本机制从该数据文件的多个副本中为用户选择一个响应时间最短的副本,称该副本为 *BestFile*。即当 *CE* 中的 *Req* 请求 *LFN* 时,使 $ResponseTime(LFN,Req)$ 最小化的 *PFN*,其中, $PFN \in |LFN|$ 。

定义 2. 删除副本 *PFN* 的代价 $PFN.cost$ 。假设副本 *PFN* 的文件大小为 $PFN.size$,当前访问频率为 $PFN.frequency$,所在节点为 *site1*,删除 *PFN* 后与 *PFN* 具有相同 *LFN* 的 *BestFile* 所在节点为 *site2*,那么

$$PFN.cost = PFN.size \times PFN.frequency / bandwidth(site1, site2),$$

其中, $bandwidth(site1, site2)$ 为节点 *site1* 与 *site2* 间传输链路的最小带宽。

定义 3. 新建 *LFN* 副本的收益 $LFN.reward$ 。假设逻辑文件 *LFN* 大小为 $LFN.size$,当前请求频率为 $LFN.frequency$,新建 *LFN* 副本所在的节点为 *site1*,新建副本前 *LFN* 的 *BestFile* 所在节点为 *site2*,那么

$$LFN.reward = LFN.size \times LFN.frequency / bandwidth(site1, site2).$$

MEDRS 策略的基本思想为:当节点 *site* 中的 *CE* 请求逻辑文件 *LFN* 时,首先判断 *site* 中是否存在 *LFN* 的副本,若存在,则返回;否则,依次判断 *site* 与 *LFN.BestFile* 所在节点路径上的每个节点是否可创建 *LFN* 的副本。另外,在判断一个节点是否可创建 *LFN* 副本时,首先判断其空闲空间(*freespace*)是否足够创建副本,若足够,则返回 True;否则,判断删除该节点中满足 $PFN.cost < LFN.reward$ 的副本后,其可用空间是否足够。MEDRS 策略的具体描述如图 3~图 5 所示。

MEDRS 策略有以下优点:首先,只有当新建副本的收益大于删除副本的代价时,才进行副本的置换,避免了副本的频繁置换;其次,当一个节点不能创建副本时,判断其邻居节点是否可以创建副本,这样可以有效降低由于文件远程访问带来的用户延迟;最后,低层节点空间不足时选择其上层节点,以适应若教育资源网格节点所在的层次越高,则其存储能力越强的特点。

```

When the CE require the LFN, the MEDRS execute
MEDRS(CE,LFN)
{
  _site=CE.site; //get the site of the CE;
  file=getbestfile(LFN,_site); //get bestfile
  site=file.site; //get the site of the file
  if (site==_site) //local site has the replication of LFN
    exit;
  shortestpath=path(_site,site); //get the shortest path between _site and site
  tempsite=_site;
  do {if (isavailablespace(tempsite,LFN)) //judge the space is available for create the replication of LFN
      {replicate(tempsite,file); //create the file's replica
        exit;} //endif
    tempsite=shortestpath.nextsite; //get the first site from the path, and set the pointer to the next
  } while (tempsite!=null)
}

```

Fig.3 MEDRS algorithm

图 3 MEDRS 算法

```

Isavailablespace(site,lfn)
{free_space=freespace(site); //get the free space of the site
if (free_space>lfn.size) //judge the space is enough
return true;
all_replicalist=allreplica(site); // sort the replica on the cost of the replica in ascending
for (i=0; i<all_replicalist.length; i++)
{if (all_replicalist[i].cost<lfn.reward)
{free_space=free_space+all_replicalist[i].size;
all_replicalist[i].pin; //if the space is enough, the pinned replica will be deleted
} //endif
else
break; //the space is not enough, endif
if (free_space>lfn.size)
return true; //the space is enough, return true
} //endifor
for (i=0; i<all_replicalist.length; i++)
all_replicalist[i].unpin; //Lack of space, set the pin with false
return false; //return the space is not enough}

```

Fig.4 Isavailablespace algorithm

图 4 Isavailablespace 算法

```

replicate(site,lfn)
{Delete the replica pinned as true on the site;
Create the replica of the lfn on the site;
}

```

Fig.5 Replicate algorithm

图 5 Replicate 算法

3.2 动态副本创建辅助策略——AEDRS

通过实际分析和应用后可以发现,在局部范围内,教育资源网的用户请求有时也存在分散性的特点,如图6所示.在这种情况下,MEDRS策略存在不足之处,我们引入AEDRS策略,间隔性地对EDRS策略进行辅助,以此解决该问题.

假设图6为市级教育资源网部分节点拓扑图, S_0 为市级管理节点, S_1 和 S_5 为县级管理节点, $S_2, S_3, S_4, S_6, S_7, S_8$ 分别为 S_1 和 S_5 辖域内的学校管理节点.如果相同大小的数据文件 lfn_1 和 lfn_2 都存储在 S_8 ,某一时间段内其文件的请求分布见表2.此时,按MEDRS策略计算 lfn_1 和 lfn_2 的副本收益, $lfn_1.reward < lfn_2.reward$.当存储空间较小时,按照MEDRS策略将为 lfn_2 创建副本,而不能为 lfn_1 创建副本,这显然是不合理的.因此,我们引入AEDRS策略,通过AEDRS策略对MEDRS策略进行辅助,以弥补MEDRS策略的不足.

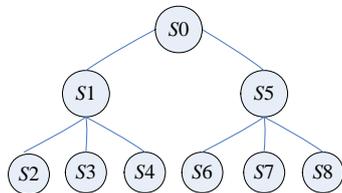


Fig.6 Site topological graph

图 6 节点拓扑图

Table 2 Request distribution of the resource

表 2 资源请求分布

Site	CEID	LFN	Frequency
S_2	CE1	lfn_1	8
S_2	CE2	lfn_1	9
S_3	CE1	lfn_2	20
S_4	CE1	lfn_1	8
S_4	CE2	lfn_1	7

AEDRS 策略的基本思想类似于 ABU 自底向上逐层聚合相同的历史记录.聚合时将历史记录中父节点相同且请求相同 LFN 的每条记录的 *frequency* 聚合起来.例如,表 1 的聚合结果为 $\langle S1,all,lfm1,32 \rangle, \langle S1,all,lfm2,20 \rangle$.另外,教育资源网格从根节点到底层节点一般不超过 6 层,分别为中心节点、省级、市级、县级、学校管理节点、底层学校资源节点.根据实际情况,相同时段内,某节点辖域内用户请求资源的相似性随着该节点所在层次的升高而降低.因此,每次聚合时并不自底向上聚合到中心节点,而是只聚合到市级节点.关于 AEDRS 策略的详细描述如图 7 所示.

```

Execute the algorithm with cycle T
AEDRS(site) //optimize the replica in the site region
{H=site.allhistory; //get the history record of the nodes in site region
roottier=site.tier; //get the tier of the site
for (tier=leafsite.tier; tier<roottier-1; tier--) //bottom-up by tier
{H=Aggregate(H); //aggregate the record
for (i=0; i<H.length; i++)
{if (H[i].site.tier==tier)
{file=getbestfile(H[i].LFN,H[i].site);
if (file.site==H[i].site) //the file's replica has existed
delete(H,H[i]); //delete the record in H
else if (isavailablespace(H[i].site,H[i].LFN))
{replicate(H[i].site,H[i].LFN); //create the replica
delete(H,H[i]);}
} //endif
} //endifor
} //endifor
}

```

Fig.7 AEDRS algorithm

图 7 AEDRS 算法

4 性能分析与测试

我们基于数据网格模拟工具 OptorSim 实现了 EDRS 策略,参考文献[5]的实验方法,利用 OptorSim 构建教育资源网格虚拟环境,采用平均作业执行时间和副本创建次数两项指标^[5],分析比较 EDRS 与 Caching-LRU 策略、Caching-LFU 策略和经济模型策略对网格系统性能的影响.实验测试时,采用如图 8 所示的网络拓扑结构,OptorSim 的参数设定见表 3.另外,为了提高实验数据的准确性,避免较大的实验误差,实验结果均取相应 10 次实验数据的平均值.

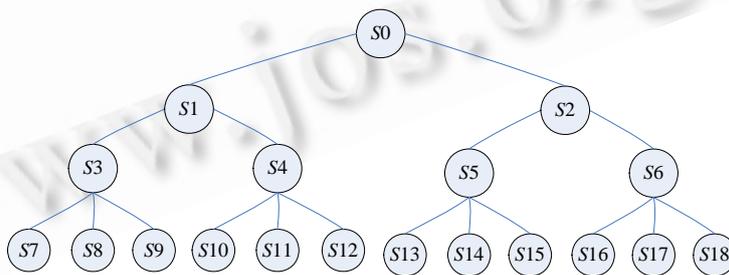


Fig.8 Topology structure of simulation experiment

图 8 模拟实验的拓扑结构

Table 3 Parameter settings of Optorsim

表 3 OptorSim 参数设定

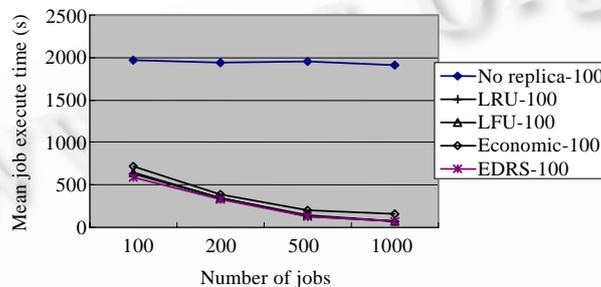
Bandwidth	Between S0 and S1, S2 is 10-50Mbps, others are 100Mbps
Size of the site	S0 is 500G; S1 and S2 are 300G; S3, S4, S5 and S6 are 200G
Initial file distribution	S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18
Access history length	100000 ms
Scheduler algorithm	Random

4.1 平均作业执行时间

平均作业执行时间是指用户提交作业的执行时间总和除以作业数量,它直接反映副本创建策略对网格系统性能的影响^[5].在相同作业量时,系统的平均作业执行时间越短,其性能越好.依据该项指标,我们分别模拟、比较并分析了作业数量、节点存储容量以及用户访问模式等因素对采用不同副本创建策略的教育资源网格系统的性能影响.

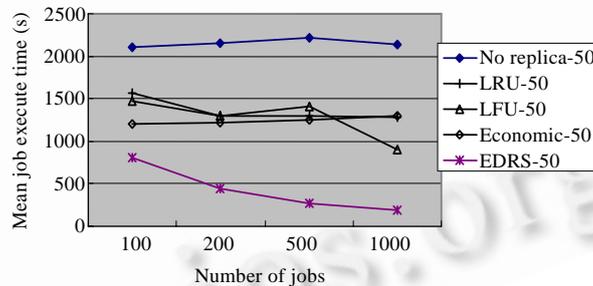
(1) 作业数量

分别设定底层节点存储容量为 100G 和 50G 进行两组实验,分析作业数量对采用不同策略的教育资源网格系统性能的影响.由图 9 可以看出,与无副本相比,各种副本放置策略都可明显降低作业的平均执行时间,这充分验证了创建副本可以提高网格系统的性能.



(a) Storage capacity is 100G

(a) 存储容量为 100G 时



(b) Storage capacity is 50G

(b) 存储容量为 50G 时

Fig.9 Relation graph of the job number and the mean job execution time

图 9 作业数量与平均作业执行时间关系图

图 9(a)是节点存储容量为 100G 时的实验结果,首先我们注意到,随着作业数的增加,4 种副本创建策略的平均作业执行时间都在减少,而且性能相差不大.这是因为当底层节点存储容量较大时,这些策略都能够在底层节点创建副本,减少了远程文件的访问次数.另外我们还可以观察到,作业数相同时,EDRS 策略性能最好,基于经济模型的副本创建策略性能最差.图 9(b)是节点存储容量为 50G 时的实验结果,我们观察到,在节点存储容量变为 50G 时,EDRS 的性能明显优于其他 3 种策略,并且只有采用 EDRS 策略时,平均作业执行时间才会随作业数的

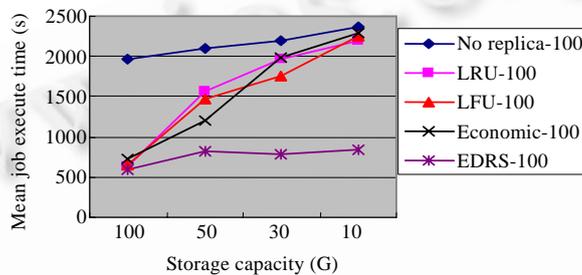
增加而减少.这是因为,当节点存储容量不足时,EDRS 能够选择邻居节点建立副本,而不是频繁地进行副本置换,从而节省了大量的远程文件访问时间.

(2) 节点存储容量

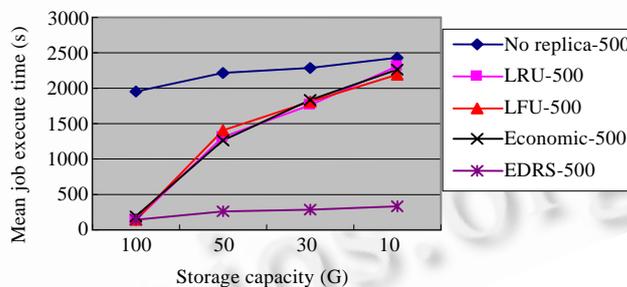
在作业数为 100 和 500 两种情况下,分析节点存储容量对采用不同策略的教育资源网格系统性能的影响.从图 10(a)和图 10(b)都可以看出,LRU,LFU 和经济模型这 3 种策略的平均作业执行时间远远大于 EDRS,当节点存储容量较小时,差异更加明显.这充分体现了在节点存储容量较小时,本文的 EDRS 策略选择邻居节点创建副本的优越性.我们还可以观察到,随着节点存储能力的减弱,LRU,LFU 和经济模型这 3 种策略的平均作业执行时间渐近无副本状态.这充分表明,这 3 种策略不适合于节点存储能力较小的教育资源网格.另外,通过图 10(a)和图 10(b)的对比我们还可以发现,即使当节点存储能力较小时,EDRS 策略的平均作业执行时间仍能随作业数的增大而减小.

(3) 访问模式

设定底层节点存储容量为 30G,作业数为 500,在分散访问和集中访问两种模式下,分析采用不同策略时系统的性能.从图 11 可以看出,提交相同的作业,分散访问的作业执行时间小于集中访问的作业执行时间.这是因为集中访问时,所有作业都集中在特定的几个节点上执行,使得底层节点存储容量不足,导致了文件的远程访问和副本的频繁置换.另外,我们测试了单独 MEDRS 策略的性能,在分散访问时,通过 MEDRS 策略和 EDRS 策略的对比可以得出,EDRS 策略的确可以进一步减少作业执行时间,提高 EDRS 的系统性能.



(a) Job num is 100
(a) 作业数为 100 时



(b) Job num is 500
(b) 作业数为 500 时

Fig.10 Relation graph of the mean job execution time and the storage capacity

图 10 存储容量与平均作业执行时间关系图

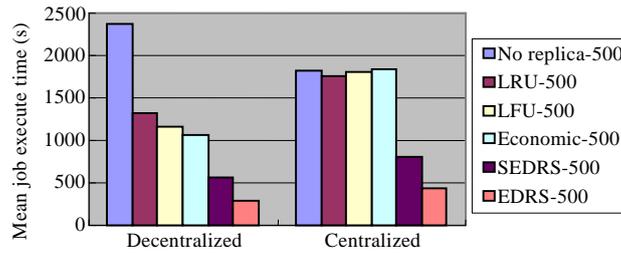
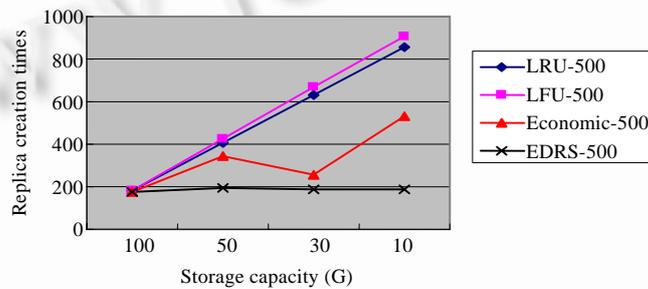


Fig.11 Relation graph of the mean job execution time and the access pattern

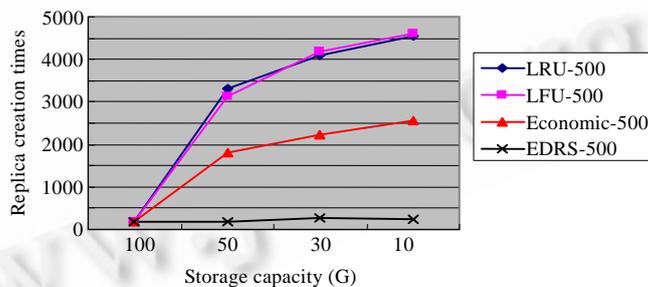
图 11 访问模式与平均作业执行时间关系图

4.2 副本创建次数

副本创建次数是指作业执行过程中创建副本的总次数.由于创建副本需要占用存储节点的 I/O 及网络带宽,所以执行相同作业时,在保证作业执行时间的情况下,副本的创建次数越少越好.我们统计了作业数为 100 和 500 两种情况下,不同策略的副本创建次数随节点存储容量的变化情况.从图 12 可以看出,执行相同作业时,随着节点存储容量的降低,LRU,LFU 和经济模型这 3 种策略创建副本次数不断增加,而 EDRS 可始终保持较低的副本创建次数.这充分说明了本文所分析的 LRU,LFU 和经济模型这 3 种策略在节点存储容量较小时容易引起副本频繁建立和删除这一结论的正确性.另外,从图 12(a)与图 12(b)的比较还可以看出,EDRS 策略的副本创建次数并不随作业数的增加而增加.这表明当底层节点存储容量不足时,选择邻居节点建立副本具有一定的优越性.



(a) Job num is 100
(a) 作业数为 100 时



(b) Job num is 500
(b) 作业数为 500 时

Fig.12 Relation graph of the replica creation times and the storage capacity

图 12 存储容量与副本创建次数关系图

4.3 EDRS策略对系统性能的影响

根据实验结果,综合作业数量、节点存储容量以及用户访问模式这 3 个因素,分析不同策略对教育资源网格系统性能的影响.另外,为了比较策略 A,B 对系统性能的改进,我们根据如下公式计算策略 B 比策略 A 性能的改进:

$$\text{性能改进}=(B-A)/A.$$

考虑到教育资源网的实际情况,采用分散式用户访问模式,分别在底层节点存储容量为 100G,50G,10G 时提交 500 次作业,统计其平均作业执行时间和副本创建次数,结果见表 4.在节点存储容量为 100G 即节点容量较大时,采用不同策略的差别不大.此时,与 LRU 相比,EDRS 策略性能仅提高 0.5%.当节点的存储容量变为 50G 时,LRU,LFU 和经济模型这 3 种策略的平均作业执行时间相差不大,但是经济模型策略的副本创建次数明显优于 LRU 和 LFU.而对于 EDRS 策略,其平均作业执行时间和副本创建次数明显优于其他 3 种策略,此时与经济模型相比,EDRS 策略的两个性能参数分别提高 78%和 89%.进一步减小节点的存储容量到 10G,则经济模型的性能开始变差,而 EDRS 策略对系统性能的改进更加明显.此时,与 LFU 策略相比,EDRS 策略的两个性能参数分别提高 84%和 95%.

从上述分析可以看出,在各种情况下,采用 EDRS 策略的教育资源网格系统性能都优于其他 3 种策略;特别是在节点存储能力较小时,EDRS 策略能够在创建较少副本的前提下获得较小的平均作业执行时间.由此可见,在节点基于一般网络环境的教育资源网格中选择 EDRS 策略是合适的.

Table 4 Comparisons of the influence of different strategies to the grid system

表 4 不同策略对网格系统性能影响的比较

Strategies	100G		50G		10G	
	Job execution time (s)	Replica creation times	Job execution time (s)	Replica creation times	Job execution time (s)	Replica creation times
No replica	1 960.1	0	2 213.5	0	2 420.2	0
LRU	135.8	183	1 302.3	3 200	2 304.3	4 559
LFU	140.6	188	1 407.6	3 210	2 191.4	4 597
Economic model	196.9	181	1 252.6	1 701	2 250.1	2 561
EDRS	135.1	179	273.2	185	339.3	240
Performance improvement of EDRS (%)	0.5	2	78	89	84	95

5 总 结

本文基于中小学教育资源的特点,提出了层次式的教育资源网格模型.通过与欧洲数据网格的对比,总结了教育资源网格与欧洲数据网格的异同.为了提高教育资源网格的服务质量,基于中小学教育资源网格的特点,重点研究了网格副本创建策略.我们发现,基于欧洲数据网格的副本创建策略,当节点存储容量较小时,容易引起副本的频繁置换,造成较大的用户延迟.本文基于网络带宽和文件大小两个参数,提出了一种副本创建策略 EDRS.实验结果分析表明,EDRS 扩展性强,适应教育资源网格的特点,能够避免副本的频繁替换,降低了用户访问延迟.同时,我们基于本文相关理论开发了中小学教育资源网格原型系统,通过模拟实验,证实了教育资源网格模型的可行性和 EDRS 策略的有效性.另外,EDRS 策略还可应用到其他普通网络环境下的网格系统中.

References:

- [1] Foster I, Kesselman C, Tuecke S. The anatomy of the grid: Enabling scalable virtual organizations. *Int'l Journal of High Performance Computing Applications*, 2001,15(3):200-222.
- [2] Chervenak A, Foster I, Kesselman C, Salisbury C, Tuecke S. The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets. *Journal of Network and Computer Applications*, 2000,23(3):187-200.
- [3] Chen L, Li SL. Dynamic self-adaptive replica location method in data grids. *Journal of Software*, 2006,17(6):1436-1447 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/1436.htm>

- [4] Ranganathan K, Foster I. Identifying dynamic replication strategies for a high-performance data grid. In: Proc. of the Int'l Grid Computing Workshop. Berlin: Springer-Verlag, 2001. 75–86. <http://www.globus.org/alliance/publications/papers/repstrat02.pdf>
- [5] Cameron DG, Carvajal-Schiaffino R. Evaluating scheduling and replica optimization strategies in OptorSim. In: Proc. of the 4th Int'l Workshop on Grid Computing (Grid 2003). Phenix: IEEE CS Press, 2003.52–59. http://www.gridpp.ac.uk/papers/optorsim_grid2003.pdf
- [6] Tang M, Lee BS, Yeo CK, Tang XY. Dynamic replication algorithms for the multi-tier data grid. Future Generation Computer Systems, 2005,21(5):775–790.
- [7] Shun HY, Wang XD, Zhou B, Jia Y, Wang HM, Zou P. The storage alliance based double-layer dynamic replica creation strategy—SADDRES. Acta Electronica Sinica, 2005,33(7):1222–1226 (in Chinese with English abstract).
- [8] Rahman RM, Barker K, Alhadj R. Replica placement in data grid: Considering utility and risk. In: Proc. of the Int'l Conf. on Information Technology: Coding and Computing (ITCC 2005). Calgary: IEEE CS Press, 2005. 354–359. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1428487
- [9] Rahman RM, Barker K, Alhadj R. Replica placement design with static optimality and dynamic maintainability. In: Proc. of the 6th IEEE Int'l Symp. on Cluster Computing and the Grid (CCGRID 2006). Washington: IEEE Computer Society, 2006. 434–437. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1630854
- [10] Gao LS, Luo JZ, Dong F, Song AB. Mesh-Based resource discovery framework in grid environment. Journal of Communication, 2006,27(11):114–118 (in Chinese with English abstract).

附中文参考文献:

- [3] 陈磊,李三立. 网格数据复本管理的动态自适应软件体系结构. 软件学报,2006,17(6):1436–1447. <http://www.jos.org.cn/1000-9825/17/1436.htm>
- [7] 孙海燕,王晓东,周斌,贾焰,王怀民,邹鹏. 基于存储联盟的双层动态副本创建策略——SADDRES. 电子学报,2005,33(7):1222–1226.
- [10] 高莉莎,罗军舟,东方,宋爱波. 基于 Mesh 结构的网格资源发现机制. 通信学报,2006,27(11):114–118.



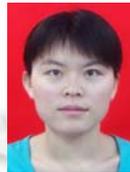
邢长明(1983—),男,山东聊城人,博士生,主要研究领域为网格计算,互连网络.



杨林(1983—),女,硕士生,主要研究领域为互连网络.



刘方爱(1962—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为互连网络,并行处理,网格计算.



厉文(1983—),女,硕士生,主要研究领域为网格计算.