数据网格虚拟机动态存储层次的研究

艾丽华,罗四维
(北京交通大学计算机与信息技术学院,北京 100044)

摘 要: 数据网格作为数据密集型计算环境,不仅需要高性能的计算资源,也需要拥有能够及时、高速访问的数据资源.数据网格中存储访问的性能是影响数据网格平台性能的关键因素.本文从网格虚拟机视角研究提高数据网格的存储访问性能,提出了针对数据网格环境的动态存储体系思想,动态 K 聚类实现网格局部性汇聚,构建数据网格动态存储体系.通过对动态存储体系性能的分析以及测试表明,具有动态存储体系的数据网格平台的任务处理性能得到了明显的改善.

关键词: 数据网格; 动态存储体系; 网格局部性汇聚; 动态 K 聚类
中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2010) 11-2680-06

Research on Dynamic Storage Hierarchy of Data Grid Virtual Machine

AI Li-hua, LUO Si-wei
(*School of Computer and Information Technology, Beijing Jiaotong University, Beijing* 100044)

Abstract: As a data-intensive computing environment, data grids require not only high performance computing resources, but also timely high speed access to the data resources, which are needed at the computational resources. The performance of the storage accesses is key to the performance of data grids platforms. The paper presents the idea of dynamic storage hierarchy appropriate for the data grid environment, and puts forward an approach of its construction by grid locality aggregation. The paper also explores the dynamic K-clustering feature of the approach, tests and analyzes the performance of the dynamic storage hierarchy. Experiment results show the explicit improvement of the overall grid tasks handle based on the dynamic storage hierarchy.

Key words: data grids; dynamic storage hierarchy; grid locality aggregation; dynamic K-clustering

1 引言

网格技术的发展将互联网的应用从共享网页信息推进到共享计算、存储以及数据资源的崭新阶段.从网格体系结构的视角,网格中间件构建了一个虚拟化的计算机,提供动态可重构的资源虚拟化功能,这里称其为网格虚拟机.网格虚拟机在互联网范围内向网格用户提供计算、存储和数据访问的能力.

数据网格作为数据密集型计算环境,其应用特点是数据动态生成量以及访问量庞大.例如,欧洲原子能研究机构 CERN 建立的大型强子对撞机 LHC 内的粒子碰撞检测器 CMS,记录加速粒子碰撞事件,每年产生几十 PB 级的数据量[1].世界各地物理学家借助数据网格平台对 CMS 产生的数据进行分析,开展相关问题探索工作.因此,研究提高数据网格环境的存储访问性能具有实际意义;与此同时,将数据网格平台视为网格虚拟机,对其存储层次特征展开研究也具有重要的学术意义.

本文从网格虚拟机视角研究提高数据网格存储访问性能,提出针对数据网格环境的动态存储体系思想,动态 K 聚类实现网格局部性汇聚,构建动态存储体系,并对构建的动态存储体系性能进行了分析和实验测试.本文利用动态聚类融网格任务管理与副本存储管理为一体,为数据网格平台内部管理机制的基础性研究奠定了基础.

2 相关工作

近年来,各国学者围绕数据网格环境的性能改善问题已经开展了若干研究.

Deng 等提出面向网格存储设备的动态可扩展存储管理结构,该体系侧重于资源虚拟化以及透明使用[2].Rahman 等采用 K 近邻原则选择最佳副本[3].该文的方法中,K 值是静态固定的.Sun 等从数据复制角度探讨了副本存储联盟概念[16].

Ranganathan 等提出分布式数据网格环境中任务执

收稿日期:2009-09-08;修回日期:2010-04-23
基金项目:国家 863 高技术研究发展计划(No.2006AA01A121);北京市教育委员会共建重点学科(No.XK100040415)

行和数据移动操作的调度问题[4].该文通过对任务调度和数据复制算法的组合测试得出结论:调度算法对系统性能有重要影响;同时考虑任务和数据的调度很重要;将任务分配到有所需数据的节点.文献[5~7]中的研究工作进一步明确了任务调度和复制机制对数据网格性能的影响.

Orlandic 针对数据网格环境中的海量存储提出两级数据聚类的层次化存储管理方法[8].从大容量和高维数据聚类角度对数据网格架构中的单一存储设备的存储布局进行研究.

本文将数据网格环境视为虚拟计算机系统.数据网格平台在处理每个网格任务时,构成了一个运行期间的虚拟机环境.在此虚拟机上执行任务的时间受到任务所需副本分布的影响,这种分布直接关系到数据网格的性能.本文提出动态存储体系思想,从数据网格虚拟机整体策略出发,对其动态存储层次展开研究.利用动态 K 聚类实现网格局部性汇聚,合理布局和重用副本文件,建立有效提高数据网格平台性能的动态存储层次.

围绕存储层次的研究早在虚拟存储器概念提出时就已经开始了.Denning 在虚拟存储器[9]一文从理论视角研究了虚拟地址到物理地址的自动分配,以及程序行为和存储管理等问题.Chow 提出了针对传统大型计算机系统的存储层次优化模型[10],指出存储体系的最优级数与系统存储容量的对数成比例,并在给定存储成本和容量的限制下,最小化存储层次的平均访问时间.Jacob 等针对指定负载局部性参数模型,推导了存储层次各级容量设置的闭合形式解[11].Du 等提出了针对机群计算平台存储层次的分析模型[12],解决在给定成本和负载类型的情况下构建、升级或扩展最优机群平台的问题.

事实上,存储层次在任何规模的计算平台上都表现出极其重要的作用.传统计算机存储层次的设计以地址变换、数据块放置和替换以及程序行为等为目标,扩大程序可访问空间,提高多道程序访问的安全性;通过建立多级存储层降低单位存储容量的成本,缩短多级存储层的平均访问时间.数据网格虚拟机是基于 Internet 的新型计算平台,该平台由网格节点动态地构成,每个网格节点包括计算元和/或存储元.本文针对网格节点本地存储元以及远程存储元所构成的两级存储层次展开研究工作.

3　数据网格虚拟机动态存储层次及其构建基础

数据网格环境中,数据对象内在的属性、存储资源的地理差异、以及网格组织的自身特点,使得数据网格虚拟机可以建立动态存储层次.动态存储层次的性能依赖于网格局部性.调度等外在算法/策略对网格局部性的汇聚加强,能够进一步改善动态存储层次的性能.

3.1　数据网格动态存储层次及特征

在数据网格环境下,一个数据源文件可能存在若干个副本文件,源文件及其副本组成了逻辑文件名 f 的物理文件集.文件分布组织表示为 $D:=(r_f^x, \leqslant T_{LR}), x=0,1,2,\cdots,K$.这里,$K$ 是文件副本数量;r_f^0 是文件源;$x\neq 0$ 时,r_f^x 是文件 f 的第 x 个副本文件.当我们不强调逻辑文件 f 所映射的具体物理标识时,可以将其物理文件简记为 r_f.$\leqslant T_{LR}$ 表示网格节点本地副本文件访问时间与远程副本文件的访问时间构成一个自反偏序关系,即访问时间满足传递性、非对称性和自反性.这里,本地(Local)文件访问比远程(Remote)文件访问耗时少,并且网格节点只从文件源或存有副本文件的节点获取本地尚未存储的数据文件,不通过中转传递.

数据网格应用中任务所需要的数据文件与其物理文件满足映像关系 $M:f\rightarrow r$,即有:

$$M(f)=r_f^x@SE_y \tag{1}$$

这里 SE_y 是物理文件 r_f^x 所在的存储元,本文用 $r_f^x@SE_y$ 表示物理文件及其位置,y 在[1..N_s]之间,N_s 是网格存储元总数.同样,当我们不强调逻辑文件 f 所映射的具体源或副本标识以及存储元位置时,可以将其简记为 $r_f@SE$.数据网格环境中的文件映像如图 1 所示.

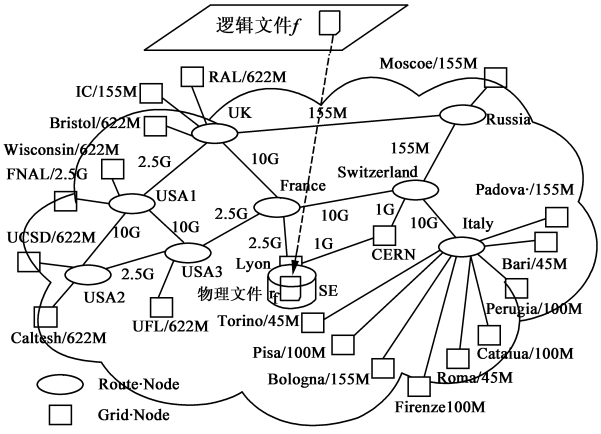

图1　数据网格环境文件映像示意

图 1 的环境中,逻辑文件与物理文件之间是一对一或一对多的关系;源文件与副本文件之间也是一对一或一对多的关系.从数据分层的角度,偏序关系、一对一关系以及一对多关系都是构成层次化问题的前提.

数据网格虚拟机有两级存储层次,分别是 H_L 本地(一级)存储层和 H_R 远程(二级)存储层.每个网格节点的存储元 SE_y 所包含的存储区域 A_{SEy} 将是两级存储层次的一个部分,即:$\forall A_{SEy}\subset(H_L\cup H_R)$.

每个存储元身份的动态变化取决于其为本地网格

任务提供了副本文件访问,还是为远程网格节点的任务请求提供副本访问.远程存储层次中包括一个特殊的节点,就是存储数据源 r_f^0 的节点,该节点不包含计算元,只作为数据源使用.

由于网格节点的动态加入和撤离、副本文件的复制与替换,使得数据网格存储层次结构可能发生变化,我们称此变化为存储层次的动态性.

在访问本地存储元发生文件缺失情况下,进行远程存储元访问.远程存储元的选择是随网格平台节点的运行状态而变化的,并非到一个固定的网格节点存储元去获取该数据文件.这使得数据网格虚拟机的存储层次具有分布性和动态性.

3.2 基于网格局部性的动态存储层次

建立数据网格动态存储层次目的是为了获取尽可能多的数据重用,使得获取数据文件所需时间尽量减小.网格局部性是数据网格虚拟机动态存储层次的构建基础.

网格局部性从任务局部性和文件局部性两个方面展现数据文件的重用.任务局部性强调将所需文件相同的任务分配到同一个节点,是主动汇聚局部性的行为.文件局部性强调数据文件在近期将被网格任务再次访问的特性,是数据文件替换时所表现的被动汇聚局部性的行为.良好的网格局部性会从时间、空间二维视角加强数据文件的汇聚,提高数据重用程度.

每个数据网格任务都需访问大量数据文件,任务所需访问的数据文件集是依赖于任务自身特点.网格任务分配策略本质上能够创建一种任务流编制模型.因此,网格任务分配编织的文件流中存在可挖掘的数据重用程度.

数据网格任务所处理的文件容量大,相比之下,网格节点存储元容量有限.在任务执行过程中,对未存储数据副本的调入会产生数据文件替换问题.因此,也需要在副本选择与替换策略方面挖掘数据文件的重用.

网格局部性对于动态存储层次的作用激励了资源调度等外在算法/策略对网格局部性汇聚的加强.

4 动态存储布局

数据网格虚拟机的存储层次是一种主动存储层次.其主动性表现在网格任务分配影响了副本文件在存储元中的布局.

4.1 K 聚类存储布局

本文利用聚类方法对网格任务流进行分析.将网格任务定义为由 d 维向量 $\boldsymbol{V}=(v_1,v_2,\cdots,v_d)$ 构成的对象空间 R^n,向量的维数与任务分配的影响因子数对应,向量的每一维分量对应所需资源的属性量化值,向量个数与网格任务数对应.聚类的目标函数如下:

$$\min \sum_{i=1}^{k}\sum_{j=1}^{m_i} \left| V_i^j - CS_i^t \right|$$

$$\text{s.t.} \sum_{i=1}^{k} m_i = n$$

这里,V_i^j 表示分配到以 CS_i 为执行点的聚类中的第 j 个网格任务,分配到 CS_i 的任务数量可达到 m_i,n 为网格任务总数;CS_i^t 是个随时间而改变的序列向量,受任务完成情况的影响,表示中心点 CS_i 当前可用资源属性,CS_i^t 的分量值代表其对应资源属性的量化值.

当 $t=0$ 时,中心点 CS_i 的计算资源尚未被任何任务占用,其所有的资源属性处于初始化状态.随着任务分配及执行的进行,中心点 CS_i 的资源属性呈现与任务执行同步的动态变化.

对于网格任务调度而言,其具体策略 P 将优先选择局部性程度高、任务队列短的网格节点,其次考虑网络传输成本开销.策略距离 D_ρ 是任务局部性程度 L_T、任务队列长度 D_Q 和网络传输成本 D_N 的函数,即有:$D_\rho(V,CS^t)=g(L_T,D_Q,D_N)$.此时,体现聚类目标的网格任务调度满足策略距离最小,见式(2);并且存储元与计算元包含于同一个网格节点 CS,见式(3).

$$\rho: V \rightarrow CS, \text{并且 } \rho(V)=\arg\min_{CS_i} D_\rho(V,CS^t) \tag{2}$$

$$\text{以及 } CS_i=(CE_i,SE_i) \tag{3}$$

任务局部性仿真群聚生物特征.这里,我们应用群聚生物外激素特性原理量化参数 L_T.例如,蜜蜂种群散发的气味对同种群的吸引,以及气味的挥发特性.网格平台初始工作时,各类任务 L_T 的量化值均为 0;执行任务 j 的计算元 CE,其 $L_T^j=1$;此时,对于该计算元上其他曾执行的任务 j',则有 $L_T^{j'}=0$.

我们利用一个非固定指派的聚类问题解决网格在线任务调度.聚类初始时,没有先验信息,也没有预先指定聚类数.聚类的结果是动态指定 K 个中心处理节点,中心点集合 $CS=\{CS_1,\cdots,CS_k\}$,进而促成 K 个中心存储元,形成 K 聚类存储布局.

4.2 动态存储层次的构建

数据网格环境中随着任务的执行,任务所需要的数据文件在网格节点中动态地被复制或者替换,构成存储层次.此过程遵循一定的映像规则和替换策略.

(1)映像规则

源文件与其副本文件或者副本文件之间存在一种映像,这种映像受网格任务及其对副本文件需求的影响.该映像规则可以具体表示为:

$$M(f \mid V)=\rho.SE.r \vee (\arg\min_{SE} D_N).r \tag{4}$$

$$\rho(V)=\left(\left|\arg\max_{CS_i}\frac{L_T}{D_Q}\right|==1\right)?\ \arg\max_{CS_i}\frac{L_T}{D_Q}:\arg\max_{CS_i\in\arg\max\frac{L_T}{D_Q}} D_N \tag{5}$$

式(4)表明将任务 V 需要的文件 f 与副本 r 映射,并且限定副本 r 位于任务调度策略 ρ 选择的存储元 SE 或者与当前存储元之间网络传输成本 D_N 最小的 SE' 之内.式(5)是一个具有判断、选择特征的形式表达式,在任务局部性程度 L_T 与任务队列长度 D_Q 的比值具有唯一最大值的情况下,选择具有该属性特征的节点,并将任务分配到此节点上;否则,在同等$\frac{L_T}{D_Q}$最大的节点范围内,选择 D_N 最小的节点.

(2)一级存储层

一级存储层由具有下述特征的存储元构成:$H_L = \{SE_i \mid (CE_i, SE_i) = CS_i \wedge \rho(V_i^j)\}$. SE_i 与执行任务的计算元 CE_i 在同一个网格节点 CS_i 上,此时可在本地访问到任务所需的副本文件,即有:$M(f \mid V) = \rho.SE.r$.

当网格任务在一级存储层中无法获得所需文件时,就到二级存储层中获取.

(3)二级存储层

二级存储层由具有下述特征的存储元构成:$H_R = \{SE_i' \mid (CE_i, SE_i) = CS_i \wedge SE_i' \neq SE_i \wedge \underset{SE_i}{\operatorname{argmin}} N_D\}$,并且,$M(f \mid V) = (\underset{SE'}{\operatorname{argmin}} D_N).r$.

二级存储层作为一级存储层的后援,保存本地没有存储的副本文件.当访问本地 SE_i 不命中时,执行任务的计算元 CE_i 与存有被访文件的 SE_i' 不在同一个网格节点上.

(4)替换规则

对于当前任务 j 访问的文件 f_i^j,如果 $f_i^j \notin H_L$,并且 H_L 的存储空间已满,就需要应用替换规则为新到来的副本文件腾出空间.本文采用基于跳-扩散随机过程的副本替换规则.

选择需要换出的文件 f_x,使其满足 $f_x = \arg\max_{f_x} D_f(Id_{f_x}, \bar{f})$,其策略距离 D_f 是当前应被替换文件的标识 Id_{f_x}与近期访问文件标识的均值$\bar{f}$ 之差.近期访问文件标识的均值$\bar{f}$受到跳-扩散随机过程控制.

在某个任务执行期间,它所访问的文件标识具有聚类相似特征;随着后继文件的访问,代表聚类中心文件标识的$\bar{f}$虽然有一定的扩散,但仍能保持该聚类的区域布局;此时替换文件属于该聚类的边缘对象.当有新任务到来时,文件标识的跳变冲出了原有的聚类区域,跳跃幅度界限为 J,即$|Id_{f_i} - \bar{f}| > J$,这使得后继访问的文件标识跳跃成为一个新的聚类布局;此时,新的聚类中心标识$\bar{f}$从 Id_{f_i}开始继续新一轮的扩散.与任务聚类不同,替换规则被动地维护本地副本文件的聚类特征.

文件标识的跳-扩散过程可综合描述为 $\mathrm{d}f_i = u_d \mathrm{d}t + \sigma_d \mathrm{d}Z(t) + J(t)\mathrm{d}P(t)$,其中 u_d 是文件标识的变化率,σ_d 是任务执行期间文件标识的波动,$Z(t)$是一个一维随机扩散过程,$J(t)$是独立、相同的随机序列,表示有新任务到来时文件标识的跳跃幅度,$P(t)$是跳跃频率为 λ 的简单泊松跳跃过程.

在换出文件 f_x 之后,就会腾出相应的存储空间 $sizeof(f_x)$.为了使得腾出的空间能够容纳 f_i^j,可能需要换出若干副本文件,这些文件腾出存储空间的总和 $sizeof(\bigcup_x \{f_x\})$应满足 $sizeof(\bigcup_x \{f_x\}) \geqslant sizeof(f_i^j)$.最后本地存储层为 $H_L' = (H_L - \{f_x\}) \cup \{f_i^j\}$.

(5)任务平均完成时间

副本文件的存储访问会影响数据网格虚拟机中网格任务的处理能力以及网络带宽的消耗.从虚拟机的角度,网格任务由处理器处理(时间为 T_P)和数据存储访问(时间为 T_M)两部分构成.网格任务平均完成时间:$ET(Job) = T_P + T_M$

$$T_M = \sum_{i=1}^{m_j} (h(f_i^j) \times t(f_i^j, H_L) + (1 - h(f_i^j)) \times t(f_i^j, H_R))$$

这里,m_j 为每个网格任务所需要访问的文件总数;任务执行过程中,数据文件 f_i^j 在本地存储元中的命中率为 $h(f_i^j)$;本地存储元的访问时间 $t(f_i^j, H_L)$受限于存储设备的性能,远程存储元的访问时间 $t(f_i^j, H_R)$受到网络和存储设备的双重影响.

在数据网格虚拟机环境,当副本存储与访问区域聚集一致时,访问缺失率就会降低,T_M 缩短;当网格任务布局合理时,处理器资源会被有效利用,T_P 缩短.网格任务布局与副本访问聚类综合考虑会有效地减小网格任务平均完成时间.

5　实验与分析

本实验利用并扩展了 OptorSim[13],模拟 CERN 的网格环境,进行了 3000 个、6 类网格任务的分配测试.每类任务的分布数量如表 1 所示.

表 1　网格任务分类统计

任务类别	JT1	JT2	JT3	JT4	JT5	JT6
任务数量	1503	579	295	316	211	96

表 1 中的 6 类任务在 CMS 测试床的 18 个计算元($CE_1 \sim CE_{18}$)上分配.每个任务都需要访问若干数据文件,文件访问模式符合单步随机漫步分布.每 GB 数据所需要的处理时间反映了网格任务的复杂程度.任务的复杂程度影响任务所需要的处理时间,进一步会影响任务队列的长度.考虑了数据处理的用时,我们在实验中对简单任务和复杂任务分别进行了测试.

图 2 示出了数据处理速度为 100ms/GB 的简单任务聚类效果.简单的任务处理用时少,文件处理速度快,任务队列短,甚至没有排队等待的情况.任务分配对计

算元呈现很强的聚类效果,非均匀特性显著.图 3 示出了数据处理速度为 100000ms/GB 的复杂任务聚类效果.复杂的任务处理用时长,文件处理速度慢,任务等待队列长.任务分配对计算元呈现均匀聚类的效果.图 2 和图 3 都表现了一个共同特性,即同一个计算元上各类任务分配数量的差异大.这种数量差异缘自任务局部性,任务倾向于同种聚类.

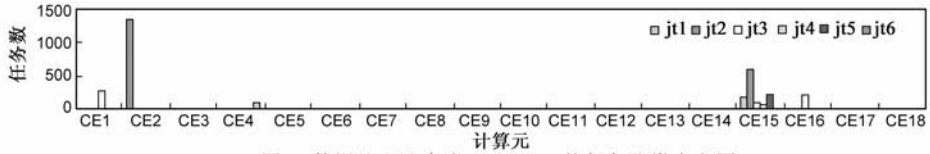

图2 数据处理速度为100ms/GB的任务聚类直方图

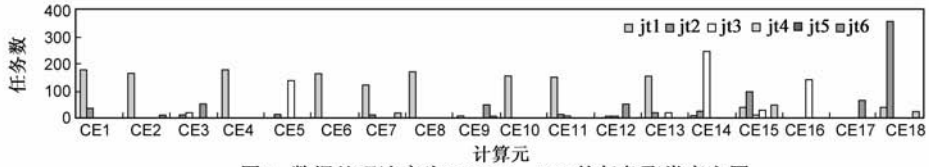

图3 数据处理速度为100000ms/GB的任务聚类直方图

数据网格任务的执行需要大量的数据文件,副本文件依赖任务的分布被复制.图 4 给出了各节点存储元内文件复制的比率,即存储元内文件复制次数与总访问次数的比值,简称副本复制率,该值反映了副本文件被重新分布的情况.这就是我们在 3.1 节提到的主动存储层次,即网格任务分配影响了副本文件在存储元中的布局.复制率高,说明副本位置变动频繁;复制率低,说明副本位置比较稳定.

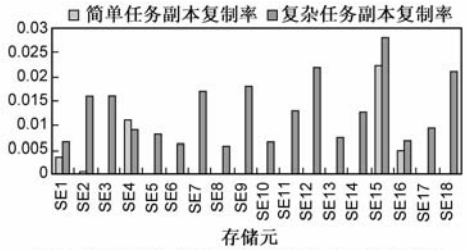

图4 简单任务及复杂任务下的副本复制率

简单任务处理在副本文件移动方面的传输量小,而复杂任务处理在此方面明显加大,这种传输量的变化影响到网络带宽的消耗.因此,提高本地访问命中率,降低远程访问率(即本地访问的缺失率)是非常重要的.图 5 和图 6 分别是简单及复杂任务下的存储节点访问命中率情况.

从简单任务到复杂任务,远程访问量增加的节点增多.这主要由于任务趋于均匀聚类而引起的副本相对频繁迁移所致.

本实验采用了动态聚类驱使的任务调度和副本替换组合策略,在简单任务及复杂任务两种情况下得到的任务平均完成时间和数据访问的平均缺失率如表 2 所示,并与另外两种策略组合即 AQcost/LRU 和 AQcost/ECO 进行了对比.LRU 是传统的近期最少使用替换规则,ECO 是近年提出的基于经济模型的副本替换规则[14].AQcost 是考虑任务队列长度和访问成本的任务调度方法,在文献[15]的实验结果中为表现最佳调度算法.

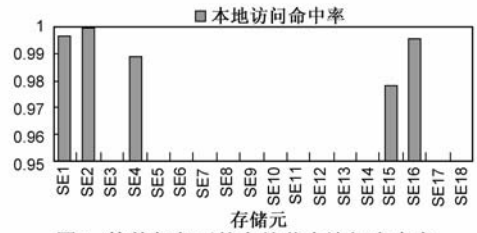

图5 简单任务下的存储节点访问命中率

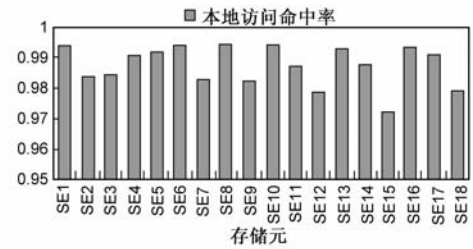

图6 复杂任务下的存储节点访问命中率

表 2 给出了改进比(Improvement Ratio)值,用来分别对比策略组合后的任务平均完成时间和副本访问的缺失率.$IR^{T}_{to\text{-}AQcost/LRU}$表示 AQcost/LRU 与本文提出的策略组合在任务平均完成时间的比值;$IR^{M}_{to\text{-}AQcost/LRU}$表示 AQcost/LRU 与本文提出的策略组合在缺失率的比值;其余类推.改进比值越大,表明被测策略组合之间的差别越大,亦即本文所提出的策略组合的优势越大.实验数据表明动态聚类驱使的任务调度和副本替换策略提高了

表 2 策略组合对比

任务复杂程度	100ms/GB	100000ms/GB
任务平均完成时间	2s	8224s
时间改进比 $IR^{T}_{to\text{-}AQcost/LRU}$	2.75	2.56
时间改进比 $IR^{T}_{to\text{-}AQcost/ECO}$	4.5	3.73
副本平均缺失率	1.06%	1.24%
缺失率改进比 $IR^{M}_{to\text{-}AQcost/LRU}$	4.05	1.68
缺失率改进比 $IR^{M}_{to\text{-}AQcost/ECO}$	1.22	1.17

数据网格平台的处理性能.

6 结论

数据访问与数据聚类是密切相关的,数据访问性能只有在数据聚类的情况下才能得到保证.本文针对网格虚拟机环境特有的动态、扩展以及重构特性,提出了以动态聚类构建数据网格虚拟机的动态存储体系.上述实验结果表明,动态聚类适应网格任务的复杂变化,保证了数据网格平台良好的任务处理性能.

参考文献:

[1] CERN openlab boosts the performance of LHC computing [EB/OL]. http://public.web.cern.ch/public/en/Spotlights-Grid_081008_en.html,2008-10-06

[2] Deng Y H, Wang F, Na H L, et al. Dynamic and scalable storage management architecture for grid oriented storage devices [J]. Parallel Computing, 2008, 34(1): 17 – 31.

[3] Rahman R M, Barker K, Alhajj R, Replica selection in grid environment: A data-mining approach [A]. Proceedings of the 2005 ACM Symposium on Applied Computing [C]. New York: ACM, 2005. 695 – 700.

[4] Ranganathan K, Foster I, Computation scheduling and data replication algorithms for data grids [EB/OL]. ftp://info.mcs.anl.gov/pub/tech_reports/reports/P1081.pdf, 2003.

[5] Chakrabarti A, Dheepak R A, Sengupta S, Integration of scheduling and replication in data grids [A]. LNCS3296 [C]. Berlin: Springer-Verlag, 2004. 375 – 385.

[6] Tang M, Lee B S, Tang X et al. The impact of data replication on job scheduling performance in the data grid [J]. Future Generation Computer Systems, 2006, 22(3): 254 – 268.

[7] Chang R S, Chang J S, Lin S Y, Job scheduling and data replication on data grids [J]. Future Generation Computer Systems, 2007, 23(7): 846 – 860.

[8] Orlandic R, Effective management of hierarchical storage using two levels of data clustering [A]. Proc 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies[C]. USA: IEEE Computer Society, 2003. 270 – 279.

[9] Denning P J, Virtual memory [J]. Computing Survey, 1970, 2(3): 153 – 189.

[10] Chow C K, On optimization of storage hierarchies [J]. IBM Journal of Research and Development, 1974, 18(3): 194 – 203.

[11] Jacob B L, Chen P M, Silverman S R, Mudge T N, An analytical model for designing memory hierarchies [J]. IEEE Transactions on Computers, 1996, 45(10): 1180 – 1194.

[12] Du X, Zhang X, Zhu Z. Memory hierarchy considerations for cost-effective cluster computing [J]. IEEE Transaction on computers, 2000, 49(9): 915 – 933.

[13] Bell W H, Cameron D G, Capozza L, et al. OptorSim: A grid simulator for studying dynamic data replication strategy [J]. International Journal of High Performance Computing Applications, 2003, 17(4): 403 – 416.

[14] Bell W H, Cameron D G, Carvajal-Schiaffino R, et al. Evaluation of an economy based file replication strategy for a data grid [A]. Proc. 3th International Symposium on Cluster Computing and the Grid [C]. USA: IEEE Computer Society, 2003. 661 – 668.

[15] Cameron D G, Carvajal-Schiaffino R, Millar A P, et al. Analysis of scheduling and replica optimisation strategies for data grids using OptorSim [J]. Journal of Grid Computing, 2004, 12(1): 57 – 69.

[16] 孙海燕,王晓东,周斌,等.基于存储联盟的双层动态副本创建策略——SADDRES [J].电子学报,2005,33(7):1222 – 1226.

Sun H Y, Wang X D, Zhou B, et al., The storage alliance based double-layer dynamic replica creation strategy-SADDRES [J]. Acta Electronica Sinica, 2005, 33(7): 1222 – 1226. (in Chinese)

作者简介:

艾丽华　女,1964 年生于哈尔滨,副教授.主要研究方向为网格计算、并行处理.
E-mail: lhai@bjtu.edu.cn

罗四维　男,1943 年生于北京,博士,博士生导师,教授.主要研究方向为网格计算、并行处理、人工神经网络、模式识别.