

文章编号:1001-9081(2005)02-0430-04

基于多物品拍卖的网格映射算法设计与模拟

刘祥瑞¹, 朱建勇², 樊孝忠¹

(1. 北京理工大学 计算机科学与工程系, 北京 100081; 2. 国家信息中心 公共技术部, 北京 100045)
(xiangruiliu@sina.com)

摘 要: 由于网格底层资源的异构、广域分布、自治等特性和网格高层的服务质量(QoS)要求,使得网格资源映射算法的设计和模拟具有相当的难度。提出了基于多物品拍卖的网格资源映射算法,解决了网格系统中相互独立任务模式的资源映射问题。同时对比了当前的网格资源映射模拟工具,阐述了利用 GridSim 建立基于多物品拍卖的资源映射算法模拟平台的主要步骤。最后模拟实验表明此算法具有优良性能。

关键词: 网格计算; 模拟; 资源映射; 多物品拍卖

中图分类号: TP393 **文献标识码:** A

Design and simulation of auction-based resource mapping algorithm in grid

LIU Xiang-rui¹, ZHU Jian-yong², FAN Xiao-zhong¹

(1. Department of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China;
2. Department of Public Technical Support, State Information Center, Beijing 100045, China)

Abstract: The design and simulation of resource mapping algorithms in grid environment is a difficult undertaking, mainly due to resource heterogeneity, geographic distribution, autonomy and the requirements of the QoS of tasks. This paper introduced auction into resource mapping and presented a resource mapping algorithm based on multi-item auction in grid. The algorithm could satisfy the requirements of the grid environment and solve the problem of the resource mapping for a set of independent tasks. And some simulation toolkits of resource mapping were analyzed. A simulation environment is established based on the Gridsim toolkit and the simulation experiments indicate the good performance of the algorithm.

Key words: grid computing; simulation; resource mapping; multi-item auction

0 引言

网格正在成为广域网上一种重要的使用模式,它的目标是融合众多的资源,建立一个功能强大的计算平台^[1]。网格中的资源包含多种类型,有计算资源、存储资源、数据资源等类型;每种类型的资源包含着大量异构的、分布的、属于不同组织的资源节点;每个资源节点又是自治的、有自身利益的、自利的主体。而网格建立的计算平台要能够提供大量可靠的计算能力、处理用户复杂的需求等等。作为连接网格底层和高层的纽带,网格资源映射算法要能够适应底层的特性、满足高层功能要求,还要能够有效配置资源。这些使得网格资源映射算法的设计和模拟具有相当的难度。现有的网格资源映射算法往往仅强调资源的有效分配和系统的全局性能优化^[2-4],没有充分处理网格底层的各种特性,而且这些算法只在简单的没有再现网格底层特性的网格环境中进行模拟,从而分析算法性能,因此也就没有表现出相应的缺陷,但这些算法往往无法适应网格复杂的环境,达不到预期的性能。如在这些算法中,映射决策所需的静态和动态信息资源往往假设通过集中或分布方式收集到信息服务中心,然而由于资源的自利性,资源显然会通过虚报自身的情况来得到更好的映射

结果,进而提高自身的收益,这就导致这些算法的性能变坏。因此对于网格中的资源映射算法不仅在设计的时候要考虑适应网格的各种要求,而且要有一个良好的环境模拟来分析其性能。

1 网格资源映射算法模拟工具

网格资源映射算法的模拟可以有多种方式来实现,如搭建实验床、单独建立模拟环境、借助模拟工具。搭建实验床方式在规模上受限于一些资源和一定的范围,很难实现模拟需要的不同的应用和资源场景,不适合算法的模拟,但由于接近实际,比较适合算法性能的简单验证。单独建立整个实验模拟环境往往需要花费很多时间,而且会出现由于考虑不周导致没有再现网格的某些特性。而借助模拟工具,不仅能够实现各种不同的应用和资源场景,而且比较节省时间。目前模拟工具主要有 Bricks^[5]、MicroGrid^[6]、SimGrid^[7]、GridSim^[8]。

Bricks 实现了在全球计算系统下不同映射算法的性能评估。它的基础模拟环境是在高性能计算机上提供科学计算函数包的以 C/S 模式运行的全球计算系统。它能够模拟资源的发现和查询、资源行为的预测、任务的虚拟处理、网络拓扑结构等。Bricks 允许开发者用 Bricks 脚本语言来配置全球计

收稿日期:2004-07-29;修订日期:2004-10-09 基金项目:国家高技术发展规划(2002AA130020)

作者简介:刘祥瑞(1976-),男,河北沧州人,博士研究生,主要研究方向:分布式计算、网格计算; 朱建勇(1972-),男,安徽泗县人,博士研究生,主要研究方向:并行计算和网格计算、计算机网络; 樊孝忠(1948-),男,河南许昌人,教授,博士生导师,主要研究方向:自然语言处理、网络信息处理。

算环境。

Microgrid 是为了研究网格资源管理而提供的一个虚拟的网格基础设施。它是建立在 Globus 平台上,因此它的基础仿真环境是利用 Globus 工具包建立起来的网格环境。它能够很好地仿真资源的行为、资源的发现、任务的处理、网络等。它的仿真结果接近实际,但需要用 Globus 构造实际的网格任务,因此适用于真实系统开发过程中对映射算法的仿真评测。

SimGrid 为模拟异构的分布式环境中的分布应用而提供一系列的核心函数。它的基础模拟环境是分布式计算平台,从简单的工作站到复杂的计算网格。它模拟了网络拓扑结构、资源的行为、任务的虚拟处理等。由于 SimGrid 提供了相当多的函数,使得模拟具有很大的灵活性,但整个过程显得比较繁琐和复杂,需要用户有一定的经验。

GridSim 为研究网格的有效资源分配技术提供一个模拟环境。GridSim 的基础模拟环境是网格计算平台,主要是市场经济的网格计算平台。它能够模拟异构而分布全球的资源、简单的网络,以及资源的查找、任务的虚拟处理等。整个框架结构比较清晰,开发者能够很容易地理解,而且由于 GridSim 采用面向对象技术,映射算法模拟开发也比较简单。

在这几种模拟工具中,GridSim 模拟工具是更具优势的。与 SimGrid 相比,GridSim 主要针对网格计算,提供了网格的各种基本功能部件,并且模拟了各个功能部件之间的基本行为,使得开发者很容易在这个模拟工具上实现算法模拟。与 MicroGrid 相比,模拟采用了虚拟时间,不受主机性能的影响,而且不需要开发实际的任务。因此 GridSim 为网格资源映射算法的研究提供了良好的基础。

2 基于多物品拍卖的网格资源映射算法

资源的映射过程是将任务分配到合适的资源上,并决定每个资源上任务的执行顺序^[9]。网格资源映射过程要适应网格底层的分布、动态、多管理域的特性,对资源要有适当的激励,以促进资源用真实的情况配合任务的分配,还要有效地配置资源、满足用户的 QoS 要求。当前网格环境下任务模式主要是一组相互独立的任务(Meta-task),本文主要分析 Meta-task 类型的资源映射。在经济学中,资源的配置问题是机制设计的中心问题,并指出拍卖在信息揭示、效率等方面的优势^[10],而且拍卖还具有简单性和分布性。结合网格的资源映射的特点,我们设计了一个拍卖过程来完成资源映射。将任务当作待招标的项目。每个资源是一个竞标者,根据自身的情况进行投标。用户代理是组织拍卖的招标者,维护整个拍卖过程,决定任务的分配。在拍卖过程中,任务的价格经过不断地调整,最后由出价最低并同时满足用户要求的资源得到相应的任务集合,同时也就完成了资源映射过程。与其他的拍卖^[11,12]相比,这个拍卖能够很好地满足网格资源映射的要求,更好地反映任务的需求,实现系统整体利益的最大化。

2.1 资源映射算法

基于多物品拍卖的资源映射算法涉及用户代理、信息服务中心、资源等多个主体,如图1所示。基于多物品拍卖的资源映射算法如下:

1) 用户代理收集和整理任务,准备资源映射即拍卖。

2) 用户代理向信息服务中心查询资源的基本信息,对资源进行初步的筛选,得到符合基本要求的资源集合 $R = \{1, 2, \dots, m\}$ 。

3) 用户代理启动任务拍卖过程。用户代理将要分配的任务集合 $I = \{1, 2, \dots, k\}$ 的任务参数和用户要求发送给集合 R 的每个资源,并通知资源招标价格和价格减少量 ε 。

4) 资源根据自身的情况,按照一定的投标规则和策略计算出要投标的任务集合及其价格,并发送给用户代理。

5) 用户代理对资源的投标进行处理,计算出临时分配。如果拍卖满足结束条件则宣布整个拍卖结束,转到6)。否则这轮招标结束,发布临时分配情况,并对每个资源的任务招标价格进行价格更新,得到新的招标价格,发送给对应的资源,开始下一轮招标,然后返回4)。

6) 用户代理根据最终分配分别通知各个资源,并将各个任务集合提交到对应的资源,资源开始对任务进行处理。资源映射结束。

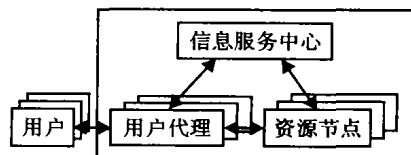


图1 基于多物品拍卖的网格资源映射简化模型

2.2 招标价格

在每一轮的开始,招标者通知每个资源 j 新的招标价格,记为 $P_j^t = \{(S, p_j^t(S)) \mid S \in I\}$,即由第 t 轮任务子集 S 和招标价格 $p_j^t(S)$ 组成。在第一轮,招标价格 P_j^1 由单个任务 i 的初始价格 $p^0(i)$ (也就是用户估价) 组成,即 $P_j^1 = \{(i, p_j^1(i)) \mid p_j^1(i) = p^0(i), i \in I, j \in R\}$ 。对于每个资源 j ,其他任务子集 S 的价格 $p_j^1(S) = \sum_{i \in S, S \in I} p^0(i)$ 。在接下来的每一轮,通知给每个资源的招标价格只是包含投标价格发生变化的任务子集及其价格的组合,即 $P_j^t = \{(S, p_j^t(S)) \mid p_j^t(S) < \min(p_j^{t-1}(S), \dots, p_j^{t-1}(S)), S \in I, j \in R\}$,其他的任务子集的价格则默认为 $\min(p_j^{t-1}(S), \dots, p_j^{t-1}(S)) \mid S \in I$ 。这样是为了任务子集的招标价格能够真实地反映竞标情况。这也表明某个任务子集的价格是与其他任务子集的价格无关。例如有任务集合 $\{a, b, c\}$,对于某个资源,在某一轮中任务集合 $\{a, b, c\}$ 的招标价格为 10,这时任务集合 $\{a, b\}$ 的招标价格不必减少到 10 或者比 10 小。

2.3 投标规则

在每一轮,资源 j 根据自身的情况提交投标集合,记为 $B_j^t = \{(S, p_{b,j}^t(S)) \mid S \in I\}$,即竞标的任务子集 S 和投标价格 $p_{b,j}^t(S)$ 组成的集合。如果投标集合中包含多个投标,则表示其中任何一个中标都可以,但不能两个或多个投标同时中标。例如对于 $B_j^t = \{(S_1, p_1), (S_2, p_2)\}$ 表示资源 j 可以以价格 p_1 接受 S_1 或者以价格 p_2 接受 S_2 ,二者只能接受一个。通常,任何任务子集 S 的投标价格 $p_{b,j}^t(S)$ 不能高于招标价格的 $p_j^t(S)$,否则投标会被招标者忽略。但有一种情况是例外的,即在每一轮一个资源 j 可以对 S 的投标价格向上浮动价格减少量 ε ,即 $p_{b,j}^t(S) \leq p_j^t(S) + \varepsilon$ 。但是一旦资源 j 采用了这种折扣,如果在接下来的招标中这个任务子集 S 的招标价格还继续减少,

就不允许 j 再投标 S 了。这个规则为了处理一个资源在上一轮没有竞标成功的任务子集的价格刚好大于它的最低价格,而在这轮中价格正好小于它的最低价格的情况。与传统拍卖算法相比,这种投标规则使得资源表达复杂利润需求,能够促进任务的有效分配。

2.4 临时分配

在每一轮结束,用户代理计算出这轮招标的临时分配,通知对应的资源。当拍卖满足结束条件的时候,最后一轮的临时分配就是最终分配,形成资源映射结果。临时分配是任务子集 S 和资源 j 的对应表,表示当前 j 在 S 上出价最低。临时分配 $\chi' = \{(j, S) \mid j \in R, S \in I\}$ 满足两个方面要求:

- 1) $\max \sum_{(j, S) \in \chi'} ((\sum_{i \in S} p^0(i)) - p_{b,j}^t(S))$
- 2) $\forall (j, S), (j', S') \in \chi', (j \neq j' \Rightarrow S \cap S' = \emptyset) \wedge (j, S) \in B_j' \wedge |\chi(j)| = 1$

这里 $|\chi(j)|$ 表示临时分配中给资源 j 任务集合的数目。临时分配将资源集合分成了两个不相交的集合:

定义 1:如果在资源的投标中的任务子集为临时分配的一部分,则称其为满意;否则为不满意。在每轮中将满意的资源集合记为 H ,不满意的资源记为 Unh 。

在拍卖过程中资源是无法得知其他资源的投标情况,只能根据自身的情况来进行投标,从而促进拍卖过程中的信息揭示。

2.5 价格更新

除了第一轮,用户代理在每轮开始的时候对任务子集的招标价格进行更新。拍卖是非匿名的,即相同任务子集的招标价格对于每个资源可能是不同的。因此价格更新是按照当前任务最低价格和价格减少量 ε ,根据每个资源所处的状态进行处理的。对于资源 j 的任务子集 S 的招标价格 $p_j^{t+1}(S)$:

$$p_j^{t+1}(S) = \begin{cases} \min(p_{b,j}^t(S), p_j^t(S)) - \varepsilon & j \in Unh \\ p_j^t(S) & j \in H \end{cases}$$

通过价格更新得到任务子集在 $t+1$ 轮的招标价格,并按照招标价格发布规则,通知对应的资源。

2.6 结束条件

当招标处于静止状态,招标就结束。下面任何一个条件都是用来判断招标是否处于静止状态。招标的结束条件是:

- 1) $\forall j \in R, j \in H$
- 2) $\forall j \in R, B_j^{t-1} = B_j^t, t \geq 2$

条件 1) 间接地表明了拍卖处于静止状态。条件 2) 要求更强一点,当连续两轮资源的投标完全相同,则直接说明招标价格和临时分配不会再发生变化了,也就是拍卖处于静止状态。

3 网格映射算法的模拟

为了详细地分析基于多物品拍卖的资源映射算法的性能,在 GridSim 模拟工具上建立了模拟平台。GridSim 工具提供了基本的模拟环境,因此整个模拟平台的建立主要包括各种结构的建立、GridSim 实体的功能扩展、调度器的开发。

结构的建立和扩展主要包括招标合同和投标合同的建立。招标合同主要描述任务、价格和其他的任务说明。投标合同主要描述资源对哪些任务子集有意向,以及每个任务子

集的价格等。

GridSim 实体的功能扩展包括 Gridlet 的扩展、资源对复杂任务的处理能力、资源的成本估价能力、资源追求自身利益的能力、资源的投标能力等。Gridlet 的扩展主要是增加了对任务的多方面描述,使用户能够更好地表达自己的需求。对资源能力扩展的目的是使资源成为一个自主的实体,模拟资源在实际网格中的情况。

调度器的开发是建立模拟平台的核心。调度器的开发主要包括标书的发布、标书的接收、临时分配的决策、任务价格的更新、拍卖的结束控制、任务的提交、任务处理结果的收集等。这样的调度器就能够组织多物品拍卖,并能够根据拍卖的结果进行资源配置。另外为了与其他资源映射算法对比,还开发了采用传统资源映射算法的调度器,扩展了采用 Buyya 的最小成本算法的调度器。

4 模拟结果

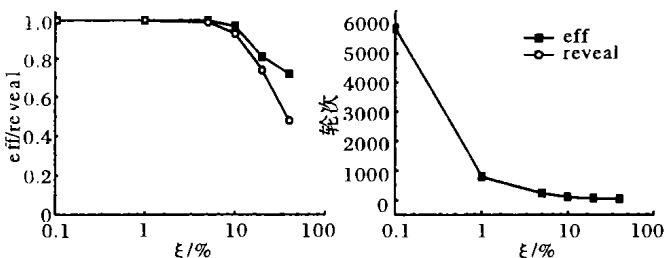
利用资源映射算法模拟平台,通过模拟实验来分析算法的性能。基于多物品拍卖的资源映射算法的性能主要表现为促进资源提供真实的情况、有效地分配资源、较好地满足用户的需求。这些分别对应为信息揭示度、有效性、用户满意度参数。信息揭示度 $reveal$ 等于所有资源的所有投标任务子集的用户收益与任务子集的利润空间之比,即:

$$reveal = \frac{\sum_{j \in R} \sum_{(S, p_{b,j}^t) \in B_j^t} (A(S) - p_{b,j}^t(S))}{\sum_{j \in R} \sum_{(S, p_{b,j}^t) \in B_j^t} (A(S) - c_j(S))}$$

有效性 eff 为拍卖后的系统收益与最优收益的比值,即:

$$eff = \frac{(\sum_{i \in I} A(i) - \sum_{S \in \chi} p(S)) + \sum_{j \in R, S \in \chi} (p(S) - c_j(S))}{(\sum_{i \in I} A(i) - \sum_{S \in \chi^*} p(S^*)) + \sum_{j \in R, S \in \chi^*} (p(S^*) - c_j(S^*))}$$

其中 $c_j(S)$ 表示资源 j 处理任务子集 S 的最低价格; $p(S)$ 表示任务子集 S 在拍卖后的成交价格; χ 表示经过多物品拍卖得到的分配结果; χ^* 表示最优分配,即这个系统收益最大时的分配。用户满意度 sa 为拍卖后用户要求得到满足的任务数目与全部任务数目的比值。



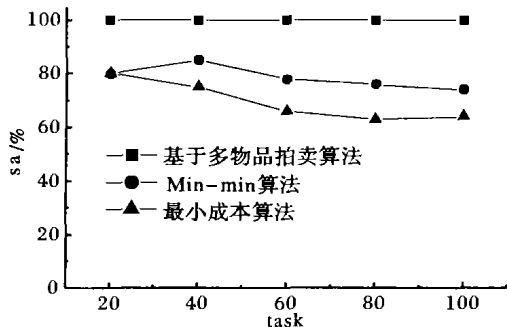
(a) 对有效性和信息揭示度的影响

(b) 对拍卖轮次的影响

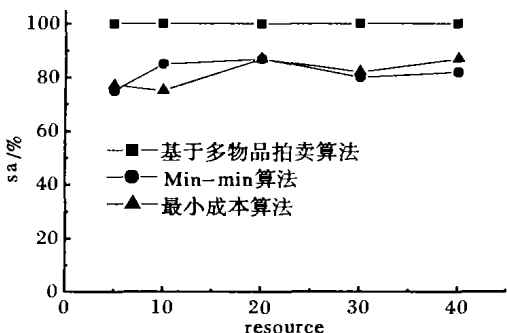
图2 价格减少量 ε 对算法性能的影响

价格减少量 ε 对算法的性能影响比较大,因此首先分析价格减少量对算法性能的影响。为了将结论在不同实验中应用,将价格减少量 ε 标准化为参数 ξ , $\xi = \varepsilon / (\Pi^* / |S^*|)$, 其中 $|S^*|$ 为最优分配时任务子集的个数, Π^* 为最优收益。图2是在10个资源和40个任务、任务的长度是在 $[0, 1000]$ 计价时间区间里均匀分布的、用户对任务没有要求、资源参数是

从 WWG^[13] 选择不同的 ξ 得到的。由于 ξ 对用户满意度 sa 没有太大的影响,因此在图2中没有列出对应的图。由图2可以得出在 ξ 比较小的时候有效性 eff 和信息揭示度 $reveal$ 都比较高,但算法的轮次较多即通信量大;在 ξ 比较大的时候通信量小,但有效性 eff 和信息揭示度 $reveal$ 就比较低。因此对于算法,需要根据对性能的要求和对通信成本的承受能力选择合适的 ξ 。一般来说 ξ 在 5% ~ 10% 的时候算法能够较好地保持高的有效性和信息揭示度,并且保持少的轮次即小的通信量。



(a) 10个资源、不同数量的任务时算法的用户满意度的对比



(b) 40个任务、不同数量的资源时算法的用户满意度的对比

图3 用户满意度对比

算法另外一个重要特点是能够全面地考虑用户的需求,实现较高的用户满意度,而传统算法缺少这点。图3对比了传统算法 Min-min 算法^[14]、Buyya 经济模型的最小成本算法^[4]和基于多物品拍卖的资源映射算法在不同数量的任务和不同数量的资源下得到的用户满意度。其中任务长度是在 [0, 1000] 计价时间区间里均匀分布;用户设定的任务完成时间则假设为任务在网格系统中的平均执行时间与等待时间之和,而任务的等待时间是 0 到所有任务平均执行时间和的区间里平均分布;任务的估价是任务在网格系统中平均成本的 1.2 倍;任务的输入和输出数据量是在 [0, 200] 计量单位的区间里均匀分布。资源参数是采自 WWG^[13]。由图3可以得出 Min-min 算法和最小成本算法的用户满意度不高。Min-min 算法主要是不能满足任务的估价要求,而最小成本算法主要是不能很好地处理资源的浮动价格,造成超过用户限定。而基于多物品拍卖的资源映射算法能够很好地考虑用户的每个需求。

5 结语

网格作为一种新的使用模式,对资源映射算法提出了新的要求。本文利用拍卖机制,建立了多物品、多轮拍卖的资源映射平台,解决了相互独立任务的资源分配。为了评价算法的性能,分析当前网格资源映射模拟工具,指出 GridSim 工具

在网格资源映射模拟方面的优势,并用 GridSim 建立了算法的模拟平台。模拟结果说明了该算法能够很好地满足各种要求,促使资源提供自身真实的情况,实现资源的有效配置,同时满足用户的个性化需求。今后进一步的工作是对有数据依赖或通信依赖的任务建立性能良好的资源映射算法,实现网格在各种环境中的应用。

参考文献:

- [1] FOSTER I. The Grid: A New Infrastructure for 21st Century Science [J]. Physics Today, 2002, 55 (2): 42-51.
- [2] MARTINO VD, MILIOTTI M. Sub optimal scheduling in a GRID using genetic algorithms [J]. Parallel Computing, 2004, 30(5-6): 553-565.
- [3] XAVIER P, LEE B-S, CAI W. A decentralized hierarchical scheduler for a grid-based clearinghouse [A]. Proceedings of the International Parallel and Distributed Processing Symposium [C]. Los Alamitos: IEEE Computer Society Press, 2003.
- [4] BUYYA R, ABRAMSON D, GIDDY J, et al. Economic models for resource management and scheduling in grid computing [J]. Concurrency and Computation Practice and Experience, 2002, 14(13): 1507-1542.
- [5] AIDA K, TAKEFUSA A, NAKADA H, et al. Performance evaluation model for scheduling in a global computing system [J]. The International Journal of High Performance Computing Applications, 2000, 14(3): 268-279.
- [6] SONG HJ, LIU X, JAKOBSEN D, et al. The MicroGrid: a Scientific Tool for Modeling Computational Grids [A]. Proceedings of the 2000 conference on Supercomputing [C]. Los Alamitos: IEEE Computer Society Press, 2000. 4-10.
- [7] LEGRAND A, MARCHAL L, CASANOVA H. Scheduling Distributed Applications: the SimGrid Simulation Framework [A]. Proceedings of the third IEEE International Symposium on Cluster Computing and the Grid [C]. Los Alamitos: IEEE Computer Society Press, 2003. 138-145.
- [8] BUYYA R, MURSHED M. GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing [J]. CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE, 2002, 14(13-15): 1175-1220.
- [9] FUDENBERG D, TIROLE J. 博弈论 [M]. 北京: 中国人民大学出版社, 2002.
- [10] BRAUN TD, SIEGEL HJ, BECK N. A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems [J]. Journal of Parallel and Distributed Computing, 2001, 61(6): 810-837.
- [11] BURGUET R. Auction Theory: A Guided Tour [J]. Investigaciones Economicas, 2000, 24(1): 3-50.
- [12] PARKES DC, UNGAR LH. An ascending-price generalized vickrey auction [EB/OL]. <http://www.eecs.harvard.edu/~parkes/pubs/iBEA.pdf>, 2002.
- [13] The Global Data-Intensive Grid Collaboration. The World-Wide Grid (WWG) [EB/OL]. <http://gridbus.cs.mu.oz.au/sc2003/list.html>, 2004.
- [14] BRAUN TD, SIEGEL HJ, BECK N. A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems [J]. Journal of Parallel and Distributed Computing, 2001, 61(6): 810-837.