

基于截止时间满意度的网格 workflow 调度算法

李 玺¹ 胡志刚^{1,2} 胡周君¹ 阎朝坤^{1,3}

¹(中南大学信息科学与工程学院 长沙 410083)
²(中南大学软件学院 长沙 410075)
³(河南大学计算机与信息工程学院 河南开封 475001)
(lixix@mail.csu.edu.cn)

Grid Workflow Scheduling Algorithm Based on Deadline Satisfaction

Li Xi¹, Hu Zhigang^{1,2}, Hu Zhoujun¹, and Yan Chaokun^{1,3}

¹(School of Information Science and Engineering, Central South University, Changsha 410083)
²(School of Software, Central South University, Changsha 410075)
³(School of Computer and Information Engineering, Henan University, Kaifeng, Henan 475001)

Abstract In grid, users usually pay more attention to the execution time of workflow than other QoS metrics. Consequently how to effectively guarantee meeting users' deadline requirements is a challenging problem for workflow scheduling in dynamic grid environment. Stochastic service model is utilized to describe dynamic processing capacity of grid resource and its dynamic workloads. The concept of deadline satisfaction degree (DSD) is defined and a corresponding calculation method for deadline satisfaction degree of workflow (DSDW) is provided. The task precedence relations represented in a DAG are converted into task execution priorities represented in numbers based on task length, and then the candidate resource for each task in the workflow is selected based on the rule of maximizing DSD. The deadline distribution is modeled as a non-linear programming problem with constraints and resolved with an interior point algorithm. A deadline satisfaction enhanced scheduling algorithm for workflow (DSESAW), which includes resource selection and overall deadline distribution, is put forward finally. The extensive simulations using real-world workflow application and grid system are made to validate this algorithm. The experimental results show that this scheduling algorithm achieves better performance than other two algorithms used in real grid system on adaptation to dynamic grid environment and user's deadline guarantee.

Key words grid; workflow; scheduling algorithm; dynamic workload; deadline

摘 要 动态网格环境中用户截止时间保障是 workflow 调度问题的一个挑战. 利用随机服务模型来描述网格资源的动态处理能力及其动态负载压力, 提出了截止时间满意度的概念和 workflow 截止时间满意度的计算方法. 将以 DAG 图形式表示的任务执行关系转换为以数值表示的任务执行优先级, 并根据最大截止时间满意度优先的思想, 确定执行 workflow 子任务的候选资源; 将 workflow 全局截止时间划分问题描述为一个约束下的非线性规划问题并通过已有方法求解该问题, 提出了一种截止时间满意度增强的 workflow 调度算法(DSESAW). 仿真实验采用实际网格应用和系统数据来验证所提出算法的性能表现, 实验结果表明新算法在网格环境的自适应性和用户截止时间保障方面优于其他两种实际网格系统中的调度算法.

关键词 网格; 工作流; 调度算法; 动态负载; 截止时间

中图法分类号 TP316.4

工作流已逐渐成为网格环境中设计大规模科学应用的典型范式^[1], 工作流执行时间通常是网格用户关心的最重要指标之一. 如何为用户提供满意度高的截止时间服务是网格工作流调度研究中的一个难点问题. 这一难点集中体现在以下两个方面: 1) Prodan 等人^[2]通过分析用户 QoS 对工作流性能的影响时指出, 工作流子任务的截止时间约束是影响工作流性能的最主要因素之一, 但用户指定的全局截止时间需求难以划分为各工作流子任务的截止时间约束; 2) Bucur 和 Epema^[3]在网格测试床 DAS-2 下作了大量的测试和比较, 其实验结果表明具有负载感知(workload-aware)的任务调度策略能获得较好的平均响应时间, 但资源负载的易变性使得在制定调度决策时难以捕获资源的动态负载信息.

针对以上两点, 本文采用随机服务模型来刻画各类网格资源的动态负载压力和动态处理能力, 首次定义了工作流截止时间满意度(deadline satisfaciton degree of workflow, DSDW)以评估一个动态负载的网格资源满足一定截止时间约束的程度, 并提出了工作流截止时间满意度的计算方法. 我们将以 DAG 图形式表示的任务执行关系转换为以数值表示的任务执行优先级, 根据最大截止时间满意度优先的思想, 确定执行工作流子任务的候选资源; 将工作流全局截止时间划分问题描述为一个约束下的非线性规划问题并通过已有方法求解该问题. 基于子任务的候选资源选择和子截止时间划分, 本文提出了一种截止时间满意度增强的工作流调度算法(deadline satisfaction enhanced scheduling algorithm for workflow, DSESAW).

1 相关工作

我们的前期工作主要集中在基于资源性能评估的独立任务调度^[4-6]上, 本文研究主要面向工作流调度算法. 网格环境下工作流任务调度研究多种多样, 其中以优化工作流执行时间为目标的突出工作有: 早期, Topcuoglu 等人^[7]通过对 DAG 图中的每个节点分配一个权重, 提出了一个异构环境下基于最早完成时间的表调度(list scheduling)算法 HEFT, 这是目前工作流调度问题中参考最多的一个经典算法. 此后, 大量学者根据不同特定领域的工作流结构

特点提出了 HEFT 的改进算法^[8-9]. 工作流调度问题中另一个经常使用的方法是分层技术, 即根据当前任务到入口/出口任务的最长路径进行正向(top level)/逆向(buttom level)分层. 苑迎春等人^[10]、Mourad 等人^[11]采用该方法, 提出了基于动态关键路径的调度算法以缩短工作流任务的执行时间.

文献[12]针对用户偏好难以有效确定的问题, 将工作流调度建模为一个扩展的多维背包问题, 采用动态规划方法求解该问题并提出了一个新的动态约束调度算法 DCA, 但 DCA 算法专用于该文提出的基于滑动约束(sliding constraint)的用户需求规范模型.

蒋哲远等人^[13]、Buyya 等人^[14-15]提出利用遗传算法、进化算法等解决费用约束下的截止时间优化问题. 由于这类算法本身寻优过程的复杂性以及实际网格环境资源数量的庞大性使得调度算法的复杂度较高.

以上研究主要是基于对描述工作流的 DAG 图的结构变换^[4-8]或采用动态规划方法^[9]、相关智能算法^[10-12]等依据资源的静态能力进行任务调度, 而系统负载的动态变化使得基于资源静态能力的调度算法无法有效保证任务的截止时间需求.

通过对资源可信性^[16]、可靠性^[17]等非功能性属性进行评估, 为工作流调度提供资源选择的依据以缩小工作流的执行时间是最近研究的一个热点. 与它们不同的是, 本文主要集中在资源动态处理能力的评估并以此进行工作流任务调度.

在调度时考虑资源动态能力的文献比较少. 著名网格项目 GRACCE^[18]中的调度器 GRACCE 主要依赖其中的 Planner 组件为上层工作流调度提供任务执行时间预测服务^[19]. 网格系统 VGrADS 中的调度器^[20]考虑多个任务竞争资源的情况, 预测任务在资源队列中的等待时间并以此选择目标资源. 这些研究提出的性能评估指标受限于特定的网格系统, 本文则屏蔽资源异构性利用统一的指标来评估资源的处理能力.

2 问题描述

设网格系统由 N 个资源节点组成, 表示为集合 (CE_1, \dots, CE_N) . 与 Bertin 等人^[21]的研究类似, 本

文采用 M/M/C 型随机服务系统来描述一个资源节点的负载和服务模型。

定义 1. 资源动态处理能力指标。

资源节点 CE_i 的动态处理能力指标表示为三元组 $\langle \lambda_i, \mu_i, c_i \rangle$, 其中 λ_i 为任务到达 CE_i 的平均时间间隔, μ_i 为 CE_i 的平均服务时间, c_i 为 CE_i 中资源的数量。

定义 2. 网格 workflow。

网格用户提交的工作流表示为二元组 $\langle D, d \rangle$, 其中, $D = (\Gamma, \Delta)$ 为以 DAG (directed acyclic graph) 图描述的工作流结构图, 集合 Γ 表示工作流中的任务 $T_i (1 \leq i \leq m)$, 即 $\Gamma = \{T_i, 1 \leq i \leq m\}$, 对应 DAG 图的节点集合; 集合 Δ 表示任务 T_i, T_j 之间的执行顺序, 对应 DAG 图的有向边集合, T_i 是 T_j 的前驱任务. d 为用户指定的工作流全局截止时间要求, 假设可以将全局截止时间参数 d 划分为各工作流子任务的截止时间参数, 并将工作流子任务的截止时间集合表示为 $d_i (1 \leq i \leq m)$ 。

定义 3. 工作流调度方案。

工作流任务到网格资源节点的一个映射, 表示为 $S: \{T_1, \dots, T_m\} \times \{CE_1, \dots, CE_N\} \rightarrow \{0, 1\}$ 。

定义 4. 工作流截止时间满意度。

对于工作流 $\langle D, d \rangle$ 给定该工作流调度方案 S , 在该方案下工作流能在截止时间 d 之前完成的概率, 表示为 $P(S, d)$ 。因此, 工作流子任务截止时间满意度为 $P(S, d_i)$ 。

3 $P(S, d)$ 计算方法

本文采用 M/M/C 型随机服务系统来描述一个资源节点的动态负载压力和处理能力模型, 并用三元组 $\langle \lambda_i, \mu_i, c_i \rangle$ 来表示资源节点 CE_i 的动态处理能力指标. Lublin^[22], Iosup^[23] 等学者对工作站集群机和实际网格系统 Grid'5000 中的资源工作负载进行监控并建模得出, 任务到达资源节点 CE_i 的时间间隔服从参数为 λ_i 的指数分布, 服务时间服从参数为 μ_i 的指数分布, 则 CE_i 的负载强度记为 $\rho_i = \lambda_i / (c_i \times \mu_i)$, 本文假设 $\rho_i \leq 1$ 。

引理 1. 工作流子任务 T_k 的截止时间满意度。

假设在调度方案 S 下, 工作流子任务 T_k 调度到资源节点 CE_i 上, 若 CE_i 的负载和服务模型用 M/M/ c_i 型队列系统描述, d_k 为子任务 T_k 的相对截

止时间要求 (相对其到达 CE_i 的时刻), 则 CE_i 能满足子任务 T_k 的截止时间要求的概率, 即工作流子任务 T_k 的截止时间满意度 $P(S, d_k)$ 为

$$Pr\{\omega \leq d_k\} = \sum_{n=0}^{c_i} \frac{\delta \times (\rho_i \times c_i)^n}{n!} + \sum_{k=1}^{c_i \times \mu_i \times d_k - 1} \delta \times \frac{\rho_i^{k+c_i} \times c_i^{c_i}}{c_i!},$$

其中, $\delta = \left[\sum_{n=1}^{c_i} \frac{(\rho_i \times c_i)^n}{n!} + \frac{(\rho_i \times c_i)^{c_i}}{c_i \times (1 - \rho_i)} \right]^{-1}$, ω 为表示任务实际完成时间的随机变量。

证明. 设随机变量 ψ 表示 CE_i 中的等待任务数目, 依据随机服务理论^[24] 可知, CE_i 中的等待任务数目为 k 的概率为

$$Pr\{\psi = k\} = \begin{cases} \delta \times \frac{\rho_i^{k+c_i} \times c_i^{c_i}}{c_i!}, & k > 0, \\ \sum_{n=0}^{c_i} \delta \times \frac{(\rho_i \times c_i)^n}{n!}, & k = 0. \end{cases} \quad (1)$$

由 M/M/ c_i 型随机服务系统的性质可知, $c_i \times \mu_i$ 为 CE_i 的服务率, 即单位时间内 CE_i 能完成的平均任务数目. 因此, 在时间段 d_k 内 CE_i 能完成的平均任务数目为 $c_i \times \mu_i \times d_k$ 。为了在截止时间 d_k 前完成当前任务, 要求 CE_i 中的等待队列长度小于或等于 $c_i \times \mu_i \times d_k - 1$ 。因此, CE_i 在截止时间 d_k 内完成任务的概率等于 CE_i 中的等待任务小于或等于 $c_i \times \mu_i \times d_k - 1$ 的概率, 即

$$Pr\{\omega \leq d_k\} = Pr\{\psi \leq c_i \times \mu_i \times d_k - 1\}, \quad (2)$$

结合式(1)和式(2)可得:

$$Pr\{\omega \leq d_k\} = Pr\{\psi \leq c_i \times \mu_i \times d_k - 1\} = \sum_{k=0}^{c_i \times \mu_i \times d_k - 1} Pr\{\psi = k\} = \sum_{n=0}^{c_i} \delta \times \frac{(\rho_i \times c_i)^n}{n!} + \sum_{k=1}^{c_i \times \mu_i \times d_k - 1} \delta \times \frac{\rho_i^{k+c_i} \times c_i^{c_i}}{c_i!}. \quad (3)$$

证毕。

推论 1. 工作流截止时间满意度。

对于工作流 $\langle D, d \rangle$ 的截止时间满意度为 $P(S, d) = \prod_{k=1}^m P(S, d_k)$ 。

由于工作流子任务 T_k 的截止时间参数 d_k 是对用户指定的全局截止时间要求 d 进行划分所得, 由引理 1 可知, 工作流子任务 T_k 的截止时间满意度 $P(S, d_k)$ 是一个以 d_k 为自变量的函数. 又根据推论 1 可知, 工作流截止时间满意度与工作流全局截止

时间 d 的划分密切相关,因此,合理的全局截止时间划分方法对用户的工作流截止时间满意度至关重要.例如假设如图 1 所示的工作流应用其截止时间划分使得仅某个子任务能够满意其截止时间要求,其余 6 个子任务对各自截止时间约束的满意度均为 0.9,则整个工作流截止时间满意度低至 0.53.

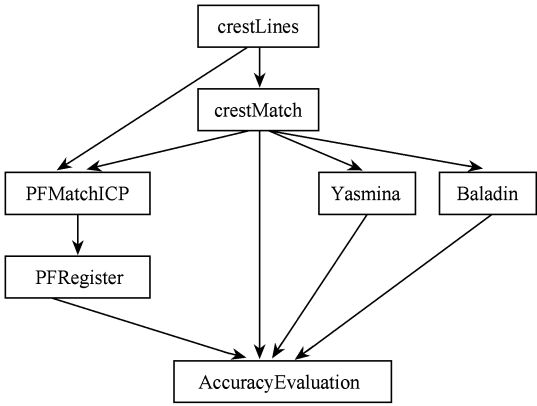


Fig. 1 Medical image analysis application.

图 1 医学图像分析工作流应用

4 工作流任务调度算法

4.1 子任务执行优先级计算方法

首先参照文献[25]中的分层技术,根据工作流子任务长度找出图 D 中的关键路径并计算关键路径上各个任务节点的优先级,将以图形式表示的任务先后执行关系转换为以数值大小表示的先后执行关系,关键路径上任务 T_i 的执行优先级如式(4)所示.然后,参照文献[15]中的同步任务(synchronization task)和简单任务(simple task)划分方法,将同步任务之间的多个分支分配相同的截止时间.若分支中存在多个子任务,各子任务的优先级计算方法如上所述.

$$rank_i = rank_j + L_i, \tag{4}$$

其中, $rank_j$ 是任务 T_i 的后继任务 T_j 的优先级, L_i 为任务 T_i 任务长度.

4.2 子任务资源选择策略

在工作流任务执行优先级确定后,对于任务 T_i 来说,无论其划分的截止时间 d_i 的多少,在所有资源中,具有最大 $P(S, d_i)$ 值的资源是唯一的.因此可将 d_i 视为常数,在所有资源中选择具有最大 $P(S, d_i)$ 值的资源作为执行任务 T_i 的候选资源.由于式(3)没有考虑多个子任务同时被调度到同一个资源节点的情况,我们在选择资源时用等待队列数组 Q

来解决这一问题,即当一个子任务被调度到一个资源节点时,其后被调度到该节点的子任务的截止时间满意度将相应地降低,任务 T_i 在资源节点 CE_j 上的截止时间满意度修正公式为

$$P_i^j = \delta_i \left[\sum_{n=0}^{c_j} \frac{(\rho_j c_j)^n}{n!} + \sum_{k=1}^{Q[j]} \frac{\rho_j^{k+c_j} c_j^{c_j}}{c_j!} \right]. \tag{5}$$

工作流子任务的资源选择策略如下:按照子任务的优先级大小降序选择资源,根据式(5)为任务 T_i 选择截止时间满意度最大的资源节点 CE_j 作为执行该任务的候选资源,当存在多个相等的最大截止时间满意度的资源时,随机选择其中一个资源作为候选资源.

4.3 工作流截止时间划分方法

根据工作流子任务资源选择策略可确定执行关键子任务的候选资源,假设执行工作流中关键子任务 $\{T_1, \dots, T_m\}$ 的候选资源分别对应为 $\{CE_{T_1}, \dots, CE_{T_m}\}$.给定工作流关键子任务的截止时间向量为 $\{d_1, \dots, d_m\}$,则各关键子任务的截止时间满意度分别对应为 $\{P_1, \dots, P_m\}$.在工作流截止时间划分方法中,我们希望找到一个最佳的截止时间向量,使得由推论 1 计算所得的工作流截止时间满意度 $P(S, d)$ 最大,该问题描述如下:

$$\begin{aligned} \max P(S, d) &= \prod_{k=1}^m P(S, d_k) \\ \text{s. t. } \sum_{i=1}^m d_k &= d, d_k > 0. \end{aligned} \tag{6}$$

该问题为可归结为一个约束条件下的非线性规划问题,其求解算法在文献[26]中有详细描述.

在确定工作流中各关键子任务的截止时间后,根据同步任务之间分支中的子任务优先级又可确定这些任务的截止时间.

4.4 工作流任务调度算法

截止时间满意度增强的工作流调度算法(deadline satisfaction enhanced scheduling algorithm for workflow, DSESAW)如图 2 所示.首先找出 DAG 图中的关键路径并根据式(4)计算关键路径中各子任务的优先级并以此对关键子任务进行排序(Setp③~Setp④);然后依次对排序后的子任务根据式(5)确定其候选资源(Setp⑥~Setp⑨),并根据文献[26]中的内点法求解各关键子任务的的最优截止时间向量 $\{d_1^*, \dots, d_m^*\}$ (Setp⑬);由于同步任务之间的一个或多个分支的截止时间与该分支并行的关键子任务截止时间相同,最后根据分支中子任务优先级计算其子截止时间(Setp⑭).

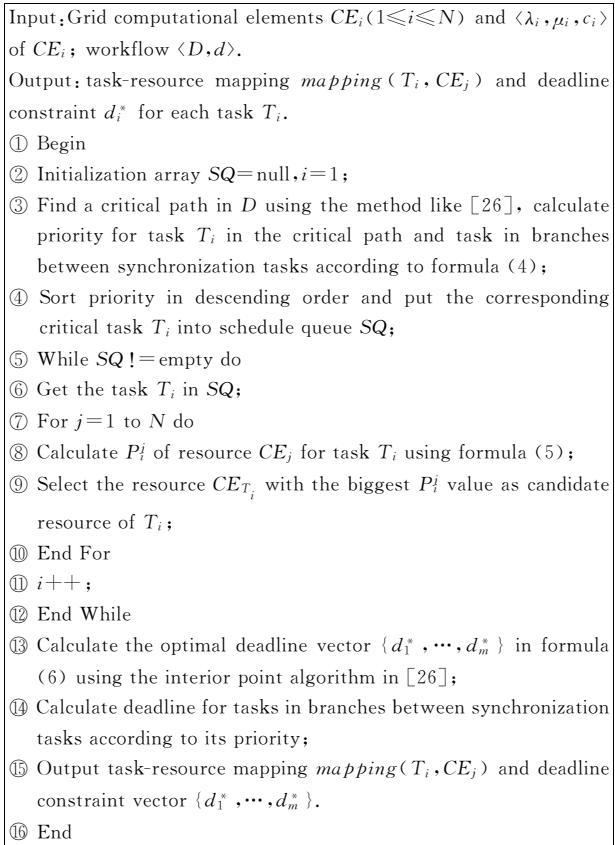


Fig. 2 DSESAW pseudo-code.

图 2 DSESAW 伪码

5 实验结果与分析

5.1 实验环境设计

基于 GridSim 5.0 网格模拟器验证提出的工作流调度算法 DSESAW 的性能. 模拟的网格系统采用法国一个实际网格系统——Grid’5000^[23]中的相关数据, 详细信息如表 1 所示:

Table 1 Simulated Grid Information
表 1 模拟网格系统的数据信息

ID	Grid Site	Processor Number	Avg. Processing Power/GHz
1	Orsay	684	2.15
2	Grenoble	68	2.5
3	Lyon	260	2.22
4	Rennes	376	2.47
5	Sophia	400	2.26
6	Bordeaux	424	2.6
7	Lille	290	2.53
8	Nancy	518	1.99
9	Toulouse	276	2.43

在表 1 中, 各节点包含的处理器数目和平均处理能力有效日期是 2010 年 4 月 1 日. 网格节点的负载模型基于文献[23]对 Grid’5000 中的工作负载档案(workloads archive)分析得出, 作业到达时间间隔参数 λ 和作业执行时间参数 μ 均为指数分布参数, 各节点的 λ, μ 参数分别在范围 $[0, 1], [1, 2]$ 内均匀分布.

用户提交的工作流为一个实际的医学图像分析应用^[27](如图 1 所示), 工作流各子任务的资源需求数量在 $[100, 180]$ 范围内均匀分布. 由于用户 QoS 参数如截止时间在现实记录中不存在, 本实验采取与文献[28]中类似的实验方法, 作业截止时间 $D = \sigma T$, 其中, T 为作业执行时间, σ 为作业截止时间紧迫因子, 在 $[1, 3]$ 范围内均匀分布.

在模拟实验中, 将本文提出的 DSESAW 算法与实际网格系统 GRACCE, VGrADS 中的调度算法在不同资源负载和不同用户截止时间要求下进行比较. 验证工作流调度问题的经典指标有:

1) 作业拒绝率(job rejection rate, JRR): 未成功完成(截止时间满意)的子任务数量与工作流任务总量的比值;

2) 作业调度长度(job schedule length, JSL): 从作业开始执行到作业完成执行消耗的时间;

3) 调度长度比(schedule length ratio, SLR): 作业调度长度 JSL 与关键路径最佳长度 E_{\min} 的比值, 即

$$SLR = JSL/E_{\min},$$

其中 E_{\min} 为 DAG 图中关键路径上的任务在各自最快资源上的执行时间之和. 该指标反映了算法调度长度与最优调度长度之间的差距.

对于每个实验场景重复 30 次实验, 取 30 次实验的平均结果作为实验结果.

5.2 实验结果与分析

1) 资源负载对工作流性能的影响

图 3 显示了不同资源负载条件下 3 种调度技术得到的工作流性能比较情况. 与 GRACCE, VGrADS 相比, 本文提出的 DSESAW 算法在 JRR 指标上平均降低 41.2%, 38.6%, 在 JSL 和 SLR 指标上平均减少 8.8%, 17.2%; 并且 DSESAW 算法最接近于最优调度长度, 平均为最优调度长度的 1.9 倍. 在资源负载压力很轻时, GRACCE 得到的 JSL 甚至稍微优于 DSESAW 算法, 这主要是因为此时资源的

能力接近于其静态能力,预测的任务执行时间接近于其实际执行时间.随着资源负载压力的增大,VGrADS 得到的工作流性能比 GRACCE 有所好转,但仍始终劣于 DSESAW 算法,这主要是因为负载压力较重时,队列等待时间构成了调度长度的一个较大比例.以上现象和数据反映:①资源负载压力对工作流的执行性能具有较大的影响;②依据随

机服务模型评估网格资源的动态能力综合考虑了任务的执行时间和队列等待时间;③提出的资源动态处理能力指标较好地刻画了网格资源的动态特征,因此提出的 DSESAW 算法对于负载动态变化的网格环境具有更好的自适应性.

2) 截止时间对工作流性能的影响

图 4 描绘了不同的作业截止时间约束下 3 种调度技术得到的作业拒绝率比较情况.在截止时间约束最为严格时,DSESAW 算法比 GRACCE,VGrADS 获得的 JRR 降低分别高达 29.6%,39.8%.随着截止时间约束的宽松,本文提出的 DSESAW 算法优势也逐渐减小.并且,截止时间的紧迫对 DSESAW,GRACCE,VGrADS 产生的 JRR 上升分别为 24.9%,28.9%,29.5%.在该实验场景中,我们还一次性提交 500 个任务,比较了同一时间段内通过不同调度技术得到的成功执行(满足截止时间约束)的任务数量(如图 5 所示).在任务开始执行 30 h,DSESAW 算法比 GRACCE,VGrADS 得到的成功调度任务数目分别提高 10.4%,29.1%.随着时间的推移,DSESAW 算法优势越来越大,GRACCE 和 VGrADS 之间的差距逐渐缩小.以上现象和数据反映:① DSESAW 算法对用户提出的截止时间要求具有更好适应性;②在多用户竞争网格资源时,DSESAW 算法能提高服务用户的数量.其二者本质原因在于 DSESAW 算法通过资源动态能力评估和全局截止时间的合理划分能获得较高的任务截止时间满意度.因此,对于由大量用户竞争和用户截止时间要求变化的网格环境中,提出 DSESAW 算法具有更好的自适应性.

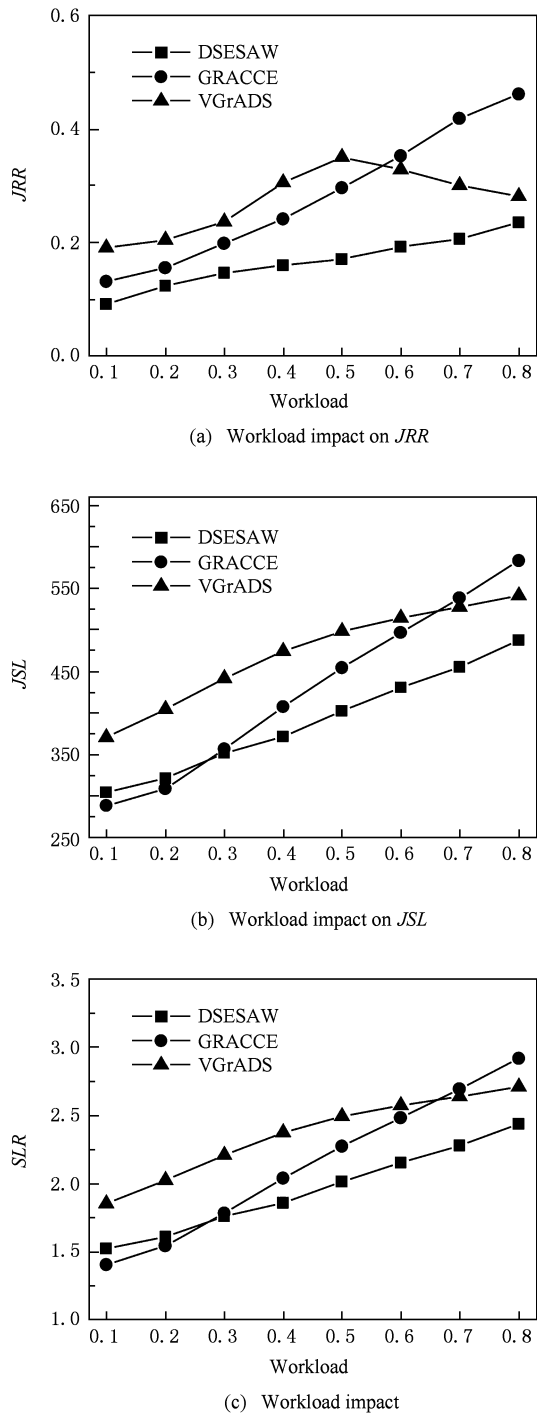


Fig. 3 Workload impact on workflow performance.
图3 资源负载对工作流性能的影响

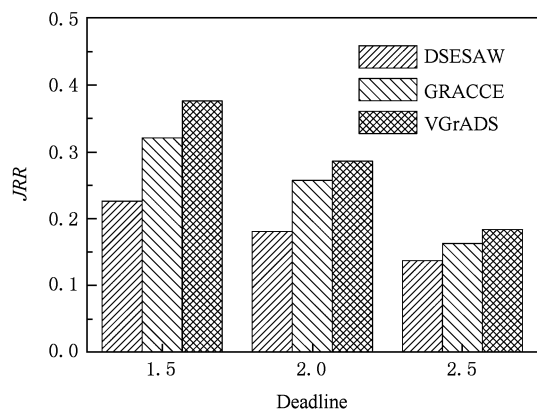


Fig. 4 Deadline impact on workflow performance.
图4 截止时间对工作流性能的影响

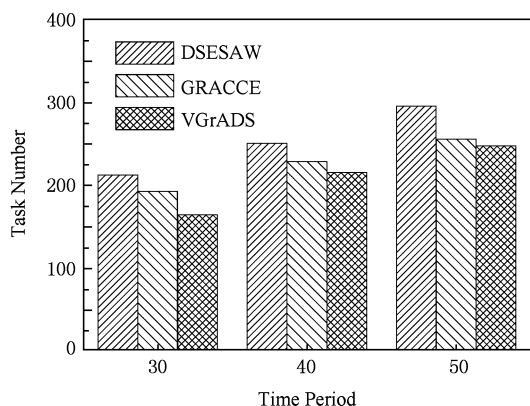


Fig. 5 Number of successfully executed tasks during a period of time.

图5 一段时间内成功执行的任务数目比较

6 结束语

用户截止时间保障是动态、异构网格环境中 workflow 调度的一个难点问题. 本文采用随机服务模型来刻画各类网格资源的动态负载压力和动态处理能力, 首次给出了 workflow 截止时间满意度 (DSDW) 的定义, 用以评估一个动态负载网格资源满意一定截止时间约束的程度, 并提出了 workflow 截止时间满意度的计算方法. 基于最大截止时间满意度优先的思想, 根据工作流子任务优先级确定了执行工作流子任务的候选资源, 将 workflow 全局截止时间划分问题描述为一个约束下的非线性规划问题并通过已有方法求解该问题, 提出了一种截止时间满意度增强的 workflow 调度算法 DSESAW. 通过大量实验验证本文算法比其他算法在网格环境的自适应性和用户截止时间保障方面的有效性. 本文对资源的动态处理能力进行了建模, 下一步工作将结合其他 QoS 指标 (如可靠性、费用等) 综合考虑资源的动态服务能力以调度 workflow 任务.

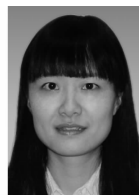
参 考 文 献

- [1] Blaha P, Schwarz K, Madsen G K H, et al. WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties [M]. Vienna, Austria: Vienna University of Technology, 2001
- [2] Prodan R, Fahringer T. Overhead analysis of scientific workflows in grid environments [J]. IEEE Trans on Parallel and Distributed Systems, 2008, 19(3): 378-393
- [3] Bucur A I D, Epema D H J. Scheduling policies for processor coallocation in multicluster system [J]. IEEE Trans on Parallel and Distributed Systems, 2007, 18(7): 958-962
- [4] Hu Zhoujun, Hu Zhigang, Li Lin. Performance-evaluation-based meta-task scheduling algorithm [J]. Journal of Xi'an Jiaotong University, 2008, 42(8): 972-976 (in Chinese)
(胡周君, 胡志刚, 李林. 一种基于性能评估的元任务调度算法[J]. 西安交通大学学报, 2008, 42(8): 972-976)
- [5] Xiao P, Hu Z -G. Deadline-guarantee-enhanced coallocation for parameter sweep application in grid [C] //Proc of IEEE Int Communications Conf (ICC 2009). Los Alamitos, CA: IEEE Computer Society, 2009: 1-5
- [6] Xiao P, Hu Z -G. A novel QoS-based coallocation model in computational grid [C] //Proc of 2008 IEEE GLOBECOM Conf. Los Alamitos, CA: IEEE Computer Society, 2008: 1562-1566
- [7] Topcuoglu H, Hariri S, Wu M. Performance effective and low-complexity task scheduling for heterogeneous computing [J]. IEEE Trans on Parallel and Distributed Systems, 2002, 13(3): 260-274
- [8] NTakpe T, Suter F, Casanova H. A comparison of scheduling approaches for mixed-parallel applications on heterogeneous platforms [C] //Proc of the 6th Int Symp on Parallel and Distributed Computing (ISPDC'07). Los Alamitos, CA: IEEE Computer Society, 2007, 35
- [9] Hunold S, Rauber T, Suter F. Scheduling dynamic workflows onto clusters of clusters using postponing [C] //Proc of the 8th IEEE Int Symp on Cluster Computing and the Grid (CCGrid'08). Los Alamitos, CA: IEEE Computer Society, 2008: 669-674
- [10] Yuan Yingchun, Li Xiaoping, Wang Qian, et al. Time optimization heuristics for scheduling budget-constrained grid workflows [J]. Journal of Computer Research and Development, 2009, 46(2): 194-201 (in Chinese)
(苑迎春, 李小平, 王茜, 等. 成本约束的网格 workflow 时间优化方法[J]. 计算机研究与发展, 2009, 46(2): 194-201)
- [11] Hakem M, Butelle F. Dynamic critical path scheduling parallel programs onto multiprocessors [C] //Proc of the 19th IEEE Int Parallel and Distributed Processing Symp (IPDPS'05). Los Alamitos, CA: IEEE Computer Society, 2005, 203
- [12] Wicczorek M, Podlipnig S, Prodan R, et al. Bi-criteria scheduling of scientific workflows for the grid [C] //Proc of the 8th IEEE Int Symp on Cluster Computing and the Grid (CCGrid'08). Los Alamitos, CA: IEEE Computer Society, 2008: 9-16
- [13] Jiang Zheyuan, Han Jianghong, Wang Zhao. An optimization model for dynamic QoS-aware Web services selection and Composition [J]. Chinese Journal of Computers, 2009, 32(5): 1014-1025 (in Chinese)
(蒋哲远, 韩江洪, 王钊. 动态的 QoS 感知 Web 服务选择和组合优化模型[J]. 计算机学报, 2009, 32(5): 1014-1025)

- [14] Yu J, Buyya R. Scheduling scientific workflow applications with deadline and budget constraints using generic algorithms [J]. Scientific Programming Journal, 2006, 14(3/4): 217-230
- [15] Yu J, Buyya R, Tham C K. Cost-based Scheduling of Scientific Workflow Applications on Utility Grids [C] //Proc of the 1st Int Conf on e-Science and Grid Computing. Los Alamitos, CA: IEEE Computer Society, 2005: 140-147
- [16] Hu Chunhua, Wu Min, Liu Guoping. QoS scheduling based on trust relationship in Web service workflow [J]. Journal of Software, 2009, 32(1): 42-53 (in Chinese)
(胡春华, 吴敏, 刘国平. Web 服务工作流中基于信任关系的 QoS 调度[J]. 软件学报, 2009, 32(1): 42-53)
- [17] Dai Y, Yang L, Zhang B. QoS-driven self-healing Web service composition based on performance prediction [J]. Journal of Computer Science and Technology, 2009, 24(2): 250-261
- [18] GRACCE: Grid Application Coordination, Collaboration and Execution [OL]. [2010-04-12]. <http://www.cs.uh.edu/~gracce>
- [19] Yan Y, Chapman B. Scientific workflow scheduling in computational grids-planning, reservation, and data/network-awareness [C] //Proc of the 8th Grid Computing Conf(Grid'07). Los Alamitos, CA: IEEE Computer Society, 2007: 18-25
- [20] Nurmi D, Mandal A, Brevik J, et al. Evaluation of a workflow scheduler using integrated performance modelling and batch queue wait time prediction [C] //Proc of the 2006 ACM/IEEE Conf on Supercomputing (SC'06). New York: ACM, 2006, 119
- [21] Bertin V, Goossens J, Jeannot E. On the distribution of sequential jobs in random brokering for heterogeneous computational grids [J]. IEEE Trans on Parallel and Distributed Systems, 2006, 17(2): 113-124
- [22] Lublin U, Feitelson D G. The workload on parallel supercomputers: Modeling the characteristics of rigid jobs [J]. Journal of Parallel Distributed Computing, 2003, 63(11): 1105-1122
- [23] Iosup A, Jan M, Sonmez O O, et al. The characteristics and the performance of groups of jobs in grids [J]. Lecture Notes on Computer Science, 2007, 46(44): 382-393
- [24] Gross D, Harris C M. Fundamentals of Queuing Theory [M]. New York: John Wiley and Sons, 1998
- [25] Du Xiaoli, Jiang Changjun, Xu Guorong, et al. A grid

DAG scheduling algorithm based on fuzzy clustering [J]. Journal of Software, 2006, 17(11): 2277-2288 (in Chinese)
(杜晓丽, 蒋昌俊, 徐国荣, 等. 一种基于模糊聚类的网格 DAG 任务图调度算法[J]. 软件学报, 2006, 17(11): 2277-2288)

- [26] Byrd R H, Hribar M E, Nocedal J. An interior point algorithm for large-scale nonlinear programming [J]. SIAM Journal on Optimization, 1999, 9(4): 877-900
- [27] Glatard T, Pennec X, Montagnat J. Performance evaluation of grid-enabled registration algorithms using bronze-standards [C] //Proc of the 9th Int Conf of Medical Image Computing and Computer-Assisted Intervention (MICCAI'06). Berlin: Springer, 2006: 152-160
- [28] Yeo C S, Buyya R. Pricing for utility-driven resource management and allocation in clusters [J]. International Journal of High Performance Computing Applications, 2007, 21(4): 405-418



Li Xi, born in 1980. PhD candidate and lecturer. Her main research interests include grid resource performance evaluation, resource allocation, and workflow scheduling.

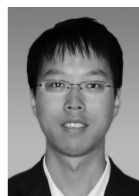


systems.

Hu Zhigang, born in 1963. Professor and PhD supervisor. Senior member of China Computer Federation. His main research interests include distributed and parallel computing, grid computing, and embedded



Hu Zhoujun, born in 1982. PhD. His main research interests include grid resource performance evaluation, resource allocation and task scheduling.



Yan Chaokun, born in 1978. PhD candidate and lecturer. His main research interests include resource reliability evaluation and workflow scheduling.