

文章编号: 1000-6788(2009)08-0102-09

## 基于博弈论健壮性增强的资源分配模型

胡周君, 胡志刚, 丁长松

(中南大学 信息科学与工程学院, 长沙 410083)

**摘要** 如何保障用户 QoS 是动态、不确定网格环境中资源分配面临的一个新问题. 针对网格资源的不可靠性特点, 提出了健壮性增强的作业执行服务的概念及其确定方法, 基于该概念提出了健壮性增强的离线资源分配模型, 通过将作业执行服务中对应的多个网格节点资源分配给作业冗余执行以增强其健壮性. 基于该模型将资源分配描述为一个非合作博弈过程, 求得了帕累托最优分配方案, 并提出了一种健壮性增强的离线资源分配策略. 实验结果表明: 该分配模型兼顾了效率和公平, 增强了作业执行的健壮性, 并较好地评估资源的信任行为.

**关键词** 计算网格; 资源分配; 博弈论; 健壮性增强; QoS

**中图分类号** TP393

**文献标志码** A

## Game-theory-based robust-enhanced model for resource allocation in computational grids

HU Zhou-jun, HU Zhi-gang, DING Chang-song

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

**Abstract** How to guarantee user's QoS for resource allocation in dynamic and uncertain Grid environment is a new challenge. Considering unreliability of grid resource, a definition called Robust-enhanced Job Execution Service (RJES) and its constructing method is proposed. A Robust-enhanced Model for Resource Off-line Allocation (RMROA) is derived from RJES. Robustness of user's job is enhanced through allocating some resource in each Grid site in RJES to the jobs redundantly. Based on RMROA, resource allocation is modeled as a non-cooperative game. A Pareto optimal allocation scheme is obtained and Robust-enhanced Off-line Allocation Policy (ROAP) is presented. We conduct extensive experiments, and the results show that RMROA not only looks after both efficiency and impartialness, but also enhances robustness of job execution and provides good mechanism for evaluating trust.

**Keywords** computational grids; resource allocation; game theory; robust-enhanced; QoS

### 1 引言

作为下一代并行与分布式计算平台, 网格计算<sup>[1]</sup> 聚集大量分散的异构资源解决大规模科学计算与工程应用问题. 这类应用由于其计算的高度复杂性通常难以由单个资源来完成, 因此需要同时对多个管理域中的

**收稿日期:** 2008-05-12

**资助项目:** 国家自然科学基金 (60673165, 60433020)

**作者简介:** 胡周君 (1982-), 男, 博士研究生, 主要研究方向为网格资源性能评估, 资源分配与任务调度; 胡志刚 (1963-), 男, 山西孝义人, 教授, 博士生导师, 主要研究领域为网络并行计算, 嵌入式系统, 网络安全.

资源进行协同分配. 如何有效地为网格应用协同分配资源已成为一个重要、开放的研究课题, 吸引了大量研究者的目光<sup>[2-3]</sup>.

由于网格资源的动态性、开放性、自治性等特点, 资源在执行任务的过程中存在诸多的不可靠因素<sup>[4]</sup> 导致资源难以保障任务的成功执行. 在资源分配时, 如何在不可靠的网格环境下保障用户 QoS 是有效资源分配的一个新的难题. 当前网格系统的资源分配主要通过资源预留机制来保障用户 QoS, 但资源预留仅仅保证资源在预留的时间段内可用, 仍然可能由于资源在该段时间内失效导致作业未能在指定的截止时间 (本文主要考虑的 QoS) 内完成; 并且过多的资源预留存在削弱系统性能、扰乱资源本地任务的正常执行等缺陷<sup>[5-6]</sup>. 因此, 如何针对网格资源的不可靠性设计 QoS 有保障的资源分配方案是一个亟待解决的问题.

我们一直致力于网格环境中任务调度与资源分配方面的研究, 我们的前期工作<sup>[7-10]</sup> 主要集中在基于资源动态可用能力评估来分配资源或调度任务. 接着前期工作, 本文则主要集中在考虑资源的不可靠特性, 如何分配资源以增强作业执行的健壮性, 从而保障用户的 QoS. 我们的长期目标是建立一套完整基于资源综合性能评估的 QoS 有保障的资源分配机制. 本文主要贡献在于:

1) 提出健壮性增强的作业执行服务 (Robust-enhanced job execution service, RJES) 概念及其确定方法, 基于该概念提出了健壮性增强的资源离线分配模型 (Robust-enhanced model for resource off-line allocation, RMROA). 本文以信任来刻画一个资源执行任务成功的可能性, RJES 包含多个信任级别不同的资源, 通过将任务调度到 RJES 中所有资源上冗余执行来保障任务的成功执行.

2) 基于健壮性增强的资源分配模型, 提出了一种基于博弈论的资源分配方案, 该方案以最大化 RJES 效用为目标, 通过 RJES 之间的博弈最终达到纳什均衡 (Nash equilibrium) 状态.

3) 通过仿真实验验证了提出的资源分配模型, 包括资源分配策略的效率与公平, 模型的健壮性增强效果, 以及资源信任评估效果等方面.

本文第 2 节介绍相关工作并与本文工作进行比较; 第 3 节建立一种健壮性增强的资源分配模型并给出了 RJES 的确定方法; 第 4 节详细讨论基于博弈论的资源分配方案; 第 5 节对本文提出的分配模型进行仿真与评估; 最后总结了全文并给出下一步工作.

## 2 相关工作

目前, 在资源分配方面存在大量的研究工作. Legion<sup>[11]</sup> 系统通过基于资源预留机制的 Enactor 组件分配资源, 存在本地任务 QoS 难以保障以及资源利用率低下等缺点. Czajkowski<sup>[3]</sup>、Foster<sup>[12]</sup> 为 Globus 系统设计了一个自动与交互式的资源分配框架. 文献 [13-14] 在资源分配时考虑资源的可用性. 其中, 文献 [13] 中的元调度器能够分配任意类型的资源, 通过与本地调度器进行协商来确定任务执行所需所有资源均处于可用状态的时间槽 (Time slot); 文献 [14] 介绍了一种可用性检查技术用于检查并更新所需资源的状态, 当所需资源均为可用状态时任务开始执行, 实验表明, 通过引入可用性检查技术可显著提高资源利用率并能减少作业响应时间. Nimrod/G<sup>[15]</sup> 系统基于计算经济思想管理和分配资源. 在该系统中, 基于截止时间和花费 QoS 提出了 3 种启发式算法<sup>[2]</sup>: 花费优化、时间优化和保守时间优化的自适应优化, 这些算法以其中一个 QoS 参数为约束条件, 以另一个 QoS 参数为优化目标, 实际上是一种单目标优化问题.

当前, 经济模型如博弈论、拍卖机制等在网格资源分配问题中得到了广泛的应用<sup>[16-20]</sup>, 这些研究大多数基于多个理性利益主体的行为产生相互影响与作用的假设, 根据资源提供者或用户的不同目标建立不同的效用函数, 并将问题描述为一个博弈过程, 通过寻求纳什均衡解达到某一目标, 提出了许多有价值的成果如资源分配方案或资源定价策略等. 文献 [21] 将资源分配模型分为合作、半合作和非合作 3 种, 通过大量实验对不同分配模型进行评估与比较, 结果表明基于代理的合作模型在任务拒绝率、资源利用率、周转时间等多个评估指标方面具有优势, 但该模型存在较大的计算复杂度.

以上研究均没有涉及到不可靠的网格环境下资源分配问题, 本文提出的资源分配模型考虑了资源的不可靠性, 并且是一种离线方式的资源分配. 提出了健壮性增强的作业执行服务概念, 通过冗余分配资源给作业以增强作业执行的成功率, 从而保障用户的 QoS; 基于该模型, 以最大化代理资源提供者效用的 *RJES* 效用

为目标, 提出了一种帕累托 (Pareto) 最优的资源分配方案.

### 3 健壮性增强的资源分配模型

#### 3.1 系统模型

如图 1 所示为一种健壮性增强的离线资源分配模型. 假设网格环境中存在  $N$  个同构的网格节点 (Site)(如一组大型机、集群或桌面机等), 表示为  $S_i(1 \leq i \leq N)$ ;  $S_i$  包含的资源 (或处理器) 数目为  $r_i$ , 并且存在一个信任属性  $t_i$ , 表示该  $S_i$  执行任务 (在本文中, 作业指用户提交的一个大型应用, 一个作业被划分成若干个子任务并调度到资源上执行) 成功的概率, 可根据  $S_i$  执行任务的历史记录获得 [22]. 健壮性增强的作业执行服务 RJES 为一组  $S_i$  的集合, 根据  $S_i$  的信任度确定 RJES 中  $S_i$  的组合方式与数目; 引入 RJES 的目的是, 当一个任务调度到 RJES <sub>$i$</sub>  中的某个节点上执行时, 同时将任务复制到该 RJES <sub>$i$</sub>  中其它所有节点上冗余执行, 以增强任务执行的成功率, 从而保证用户的截止时间 QoS. 元调度器 (Meta-scheduler) 负责对一段时间内多个用户提交的作业协同分配资源, 并将任务调度到 RJES 中所有节点上冗余执行.

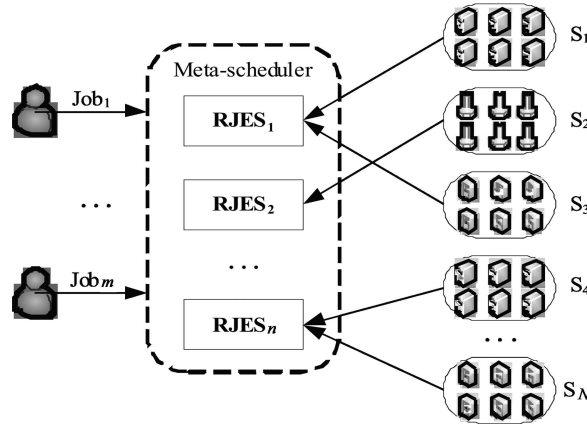


图 1 健壮性增强的离线资源分配模型

#### 3.2 RJES 确定方法

冗余调度技术通常面临的一个问题是如何确定副本的数量, 在本文体现为如何确定 RJES 中网格节点的组合成分与数目. 本文以信任刻画一个资源执行任务成功的概率, 通过将任务调度到多个节点上冗余执行确保任务的成功执行, 假设某个 RJES 包含  $s$  个网格节点  $S_i(1 \leq i \leq s)$ , 则该 RJES 的信任度 (即执行任务成功率) 与 RJES 中资源数目分别为:

$$t_{RJES} = 1 - \prod_{i=1}^s (1 - t_i) \quad (1)$$

$$X_i = \min\{r_1, r_2, \dots, r_s\} \quad (2)$$

输入:  $S_i(1 \leq i \leq N)$  及其信任值  $t_i$ ;

输出:  $RJES_k$  及其组成成分  $S_j$ .

1. 初始化  $k = 1$ ;
2. 将  $S_i$  按其信任值  $t_i$  降序排列, 并存放到一个队列  $Q$  中;
3. while ( $Q$  is nonempty)
4. 将队列首元素  $Q_{first}$  指定的网格节点  $S$  添加到  $RJES_k$  中, 并删除该  $S$ ;
5. while ( $t_{RJES_k} < t_{th}$ )
6. 将队列尾元素  $Q_{last}$  指定的网格节点  $S$  添加到  $RJES_k$  中, 并删除该  $S$ ;
7.  $k++$ ;

图 2 RDH 算法

向 RJES 中添加网格节点直至满足条件  $t_{RJES_k} > t_{th}$  即可,  $t_{th}$  表示 RJES 信任度阈值, 可根据经验得出, 含义是当 RJES 信任度大于该阈值时, 任务在该 RJES 中的网格节点上冗余执行能以较大概率成功执行

完毕. 因此, RJES 中网格节点的确定问题可简化为一个 0/1 背包问题. 由于 0/1 背包问题的 NP 难特性<sup>[23]</sup>, 本文设计了一个 RJES 确定启发式算法(RJES determination heuristic, RDH)(如图 2) 来确定 RJES 中的网格节点.

该算法具有如下特点: 1) 对于新加入的节点或任务执行历史记录较少的节点, 其低信任度难以刻画该节点真正行为, 通过将高信任节点与低信任节点放入同一个 RJES 中冗余执行任务, 不但可以增强任务执行的成功率, 而且能给行为好或有此趋向的节点提供执行任务的机会, 使其提高信任值, 从而提高对该资源信任评估的准确性; 2) 算法具有较低的复杂度  $O(n)$ , 但由于节点信任度的动态变化, 该算法需周期性运行.

## 4 基于健壮性增强系统模型的资源分配方案

### 4.1 资源分配的非合作博弈

对于某段时间内到达的  $m$  个作业, 资源提供者提供资源给作业执行时能获得一定的效用 (Utility), 通过 RJES 代理网格节点中资源提供者的效用, 对文献 [24] 中效用函数进行改进, 本文中 RJES 效用函数表示为:

$$u(x) = \omega \log(\varepsilon x), \quad (3)$$

其中,  $u$  表示  $RJES_i$  中每个网格节点均分配  $x$  数量 (处理器数目或处理器能力) 的资源给某个作业时所得到的效用;  $\omega$  表示不同紧迫程度的作业对效用的影响,  $\omega = D/(D - T_{\text{exp}})$ ,  $D$  表示作业的截止时间,  $T_{\text{exp}}$  表示该作业期望完成时间;  $\varepsilon$  为常量, 表示不同 RJES 对效用的偏好.

RJES 资源分配的非合作博弈描述如下:

· 局中人 (Player):  $n$  个 RJES;

· 策略 (Strategy): 局中人  $RJES_i$  采取的策略是分配给作业  $Job_j$  的资源数量  $x_{ij} (1 \leq i \leq n)$ , 即该  $RJES_i$  中所有网格节点均分配  $x_{ij}$  数量的资源给  $Job_j$  (当作业开始执行时, 将  $Job_j$  中任务同时调度到  $RJES_i$  中所有网格节点上冗余执行, 并占用每个节点  $x_{ij}$  数量的资源), 并且所有 RJES 分配给  $Job_j$  的资源数量同该作业的资源需求数量  $J_j (1 \leq j \leq m)$  相等, 即:

$$\sum_{i=1}^n x_{ij} = J_j \quad (4)$$

当局中人策略  $x_{ij} (2 \leq i \leq n)$  已知时, 根据式 (4) 可得  $RJES_1$  策略为

$$x_{1j} = J_j - \sum_{i=2}^n x_{ij} \quad (5)$$

局中人策略的一个限制条件是,  $RJES_i$  分配给  $m$  个作业的资源数量之和小于该  $RJES_i$  中资源数量, 即:

$$\sum_{j=1}^m x_{ij} \leq X_i \quad (6)$$

· 收益 (Revenue): 分配资源给  $Job_j$  获得的收益根据分配给该作业的资源数量  $x_{ij} (2 \leq i \leq n)$  及其效用函数计算可得. 因此, 分配资源给  $Job_j$  获得的收益为:

$$\begin{aligned} R_j &= \omega_j \log(\varepsilon_1 x_{1j}) + \omega_j \log(\varepsilon_2 x_{2j}) + \cdots + \omega_j \log(\varepsilon_n x_{nj}) \\ &= \omega_j \log \left( \varepsilon_1 \left( J_j - \sum_{i=2}^n x_{ij} \right) \right) + \omega_j \log(\varepsilon_2 x_{2j}) + \cdots + \omega_j \log(\varepsilon_n x_{nj}) \end{aligned} \quad (7)$$

$n$  个 RJES 根据各自的收益及其它 RJES 的策略不断地进行调整, 最终达到纳什均衡状态. 达到纳什均衡时, 在其它 RJES 策略不变时, 某个 RJES 改变其策略时不能增大总收益<sup>[25]</sup>, 此时的分配策略  $(\hat{x}_{2j}, \hat{x}_{3j}, \cdots, \hat{x}_{nj})$  达到帕累托最优. 即:

$$\begin{aligned} R((\hat{x}_{21}, \hat{x}_{31}, \cdots, \hat{x}_{n1}), \cdots, (\hat{x}_{2j}, \hat{x}_{3j}, \cdots, \hat{x}_{nj}), \cdots, (\hat{x}_{2m}, \hat{x}_{3m}, \cdots, \hat{x}_{nm})) &\geq \\ R((\hat{x}_{21}, \hat{x}_{31}, \cdots, \hat{x}_{n1}), \cdots, (x_{2j}, x_{3j}, \cdots, x_{nj}), \cdots, (\hat{x}_{2m}, \hat{x}_{3m}, \cdots, \hat{x}_{nm})) & \\ \forall (x_{2j}, x_{3j}, \cdots, x_{nj}) & \end{aligned} \quad (8)$$

## 4.2 最优资源分配策略

**引理 1** 分配资源给  $Job_j$  获得的收益  $R_j$  存在最大值.

**证明** 将式 (7) 中  $R_j$  对  $x_{ij}(2 \leq i \leq n)$  分别求一阶偏导和二阶偏导, 即:

$$\frac{\partial R_j}{\partial x_{ij}} = -\frac{\omega_j}{J_j - \sum_{k=2}^n x_{kj}} + \frac{\omega_j}{x_{ij}} \quad (9)$$

$$\frac{\partial^2 R_j}{\partial x_{ij}^2} = -\frac{\omega_j}{\left(J_j - \sum_{k=2}^n x_{kj}\right)^2} - \frac{\omega_j}{x_{ij}^2} \quad (10)$$

故,  $\frac{\partial R_j}{\partial x_{ij}} \geq 0$  或  $\frac{\partial R_j}{\partial x_{ij}} < 0$ ,  $\frac{\partial^2 R_j}{\partial x_{ij}^2} < 0$ . 因此,  $R_j$  是  $n-1$  维空间中的一个凹函数. 同时,  $R_j$  在其定义域上连续, 故  $R_j$  存在最大值.

在一定限制条件下 (式 (6)), 求取分配资源给某个作业时获得的收益  $R_j$  (式 (7)) 达到最大时  $RJES_i$  的分配策略  $(x_{2j}, x_{3j}, \dots, x_{nj})$ , 可以用拉格朗日法求之.

令拉格朗日函数

$$L_j = R_j + \lambda \left( X_i - \sum_{k=1}^m x_{ik} \right) \quad (11)$$

将式 (7) 代入式 (11) 得,

$$L_j = \omega_j \log \left( \varepsilon_1 \left( J_j - \sum_{i=2}^n x_{ij} \right) \right) + \omega_j \log(\varepsilon_2 x_{2j}) + \dots + \omega_j \log(\varepsilon_n x_{nj}) + \lambda \left( X_i - \sum_{k=1}^m x_{ik} \right) \quad (12)$$

将式 (12) 对  $x_{ij}(2 \leq i \leq n)$  (可将  $j$  视为常量) 求偏导, 即:

$$\frac{\partial L_j}{\partial x_{ij}} = 0 \implies -\frac{\omega_j}{J_j - \sum_{k=2}^n x_{kj}} + \frac{\omega_j}{x_{ij}} - \lambda = 0 \quad (13)$$

将式 (5) 代入式 (13), 解得:

$$x_{ij} = \frac{\omega_j x_{1j}}{\omega_j + \lambda x_{1j}}, \quad 2 \leq i \leq n \quad (14)$$

将式 (14) 代入式 (4), 得:

$\lambda x_{1j}^2 + (n\omega_j - \lambda J_j)x_{1j} - \omega_j J_j = 0$ , 解方程得:

$$x_{1j} = (\lambda J_j - n\omega_j + \delta)/2\lambda \quad (15)$$

其中,  $\delta = \sqrt{(n\omega_j - \lambda J_j)^2 + 4\lambda\omega_j J_j}$ .

将式 (15) 代入式 (6), 即:

$$\sum_{j=1}^m (\lambda J_j - n\omega_j + \delta)/2\lambda \leq X_1 \quad (16)$$

解不等式, 令

$$\lambda \leq f(\omega_j, J_j, X_1) \quad (17)$$

**引理 2** 当  $\lambda = \lambda_{\max} = f(\omega_j, J_j, X_1)$  时,  $R_j$  最大.

**证明** 由式 (14)、(15) 可知,  $x_{1j}$  为  $\lambda$  的单调递减函数,  $x_{ij}$  为  $\lambda$  的单调递增函数.

将式 (5) 代入式 (9) 得,  $\frac{\partial R_j}{\partial x_{ij}} = -\frac{\omega_j}{x_{1j}} + \frac{\omega_j}{x_{ij}}$ .

1) 当  $x_{1j} < x_{ij}(2 \leq i \leq n)$  时,  $\frac{\partial R_j}{\partial x_{ij}} < 0$ , 则  $R_j$  为  $x_{ij}$  的单调递减函数, 因此, 当  $x_{ij} = \min x_{1j}$  时,  $R_j$  最大. 又由于  $x_{1j}$  为  $\lambda$  的单调递减函数, 当  $\lambda$  取最大值时, 即  $\lambda = f(\omega_j, J_j, X_i)$  时,  $x_{1j}$  最小.

2) 当  $x_{1j} > x_{ij}$  时,  $\frac{\partial R_j}{\partial x_{ij}} > 0$ , 则  $R_j$  为  $x_{ij}$  的单调递增函数, 又由于  $x_{ij}$  为  $\lambda$  的单调递增函数, 则  $R_j$  为  $\lambda$  的单调递增函数, 因此,  $\lambda = f(\omega_j, J_j, X_i)$  时,  $R_j$  最大.

此时, RJES 的分配策略如下:

$$x_{1j} = (\lambda_{\max} J_j - n\omega_j + \delta)/2\lambda_{\max}, x_{ij} = \frac{\omega_j x_{1j}}{\omega_j + \lambda_{\max} x_{1j}} (2 \leq i \leq n, 1 \leq j \leq m) \quad (18)$$

则由式 (18) 确定的资源分配数量为帕累托最优. 基于该纳什均衡点, 提出一种健壮性增强模型的离线分配策略 (Robust-enhanced off-line allocation policy, ROAP). 该策略对一段时间内到达的  $m$  个作业同时分配多个资源, 根据式 (18) 确定分配给每个作业的每个 RJES 中资源数量, 并冗余分配该 RJES 中所有网格节点中该数量的资源给该作业.

## 5 性能评估

### 5.1 仿真实验设计

基于 GridSim<sup>[26]</sup> 网格模拟器验证提出的资源分配模型的性能. 模拟的网格系统采用法国一个真实网格系统——Grid'5000<sup>[27]</sup> 中的相关数据, 详细信息如表 1 所示, 其中, 各节点包含的处理器数目有效日期是 2006 年 12 月 12 日, 信任度根据文献 [28] 对 Grid'5000 中的资源可用性、可靠性追踪数据计算得出. 网格节点的效用偏好  $\varepsilon$  与该节点处理器能力成正比,  $\varepsilon = k \cdot pw$ ,  $pw$  表示该节点处理器能力,  $k$  为比例系数, 在 [1, 2] 范围内均匀分布, RJES 效用偏好  $\varepsilon$  为该 RJES 中所有网格节点效用偏好的均值. RJES 信任度阈值  $t_{th} = 0.9$ .

用户提交的作业根据文献 [29] 对 Grid'5000 中的工作负载档案 (Workloads archive) 分析得出结论, 作业执行时间服从参数为  $\rho$  的指数分布, 实验中  $\rho = 0.001$ , 作业在某天内到达数目  $m = 10$ , 作业的资源需求数量  $J_j (1 \leq j \leq m)$  在 [120,170] 范围内均匀分布, 资源供大于求. 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 [30] 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,  $\sigma$  为作业截止时间紧迫因子, 在 [1,3] 范围内均匀分布.

表 1 模拟网格系统的数据信息

Grid site	No. processors	Processing power (/GHZ)	Trust degree
S1	128	2.2	0.57
S2	128	2.2	0.71
S3	198	2.2	0.72
S4	64	3.2	0.59
S5	196	1.6	0.75
S6	114	2.4	0.91
S7	210	1.6	0.87
S8	112	2.4	0.82
S9	112	2.4	0.83
S10	139	2.2	0.92
S11	206	1.6	0.90
S12	64	3.0	0.73
S13	92	2.6	0.88
S14	684	1.4	0.91
S15	146	2.0	0.75

模拟实验中, 将本文提出的 ROAP 分配策略与以下资源分配策略进行比较. 1) 基于能力的随机策略 (Capability-based random policy, CRP): 选择一个网格节点的概率与该节点处理器能力成正比; 2) 节点最小策略 (Site minimized policy, SMP): 调度器选择一个节点集合, 使得集合中节点数目最小. 主要从以下 2 个性能指标对上述 3 种策略进行比较: RJES 收益和作业拒绝率 (Job rejection ratio, JRR). 此外, 模型中的 RDH 算法增加了信任度低的节点的执行任务机会, 因此, 对于某些执行任务历史记录少或有向行为好的趋势发展的节点, 需要验证这些资源信任增长情况, 从而准确地评估一个资源的信任度. 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) 收益比较

图 3 描绘了 ROAP 策略下各 RJES 收益及其能力 (点旁数字) 比较情况, 可以看出, RJES 能力越强, 收益越高. 图 4 描绘了 3 种分配策略下产生收益的比较情况. 从图 4 可以看出, 在大多数情况下, ROAP 分配策略产生的收益均大于 CRP 和 SMP 策略产生的收益. 因为 CRP 策略经常选择能力大的节点来执行任务, 对于能力小的节点被选中的概率小, 从而得到的收益较少; SMP 策略总是选择处理器数目多的节点来执行任务, 从而使得处理器数目少的节点收益也少. 而 ROAP 策略中, 所有 RJES 根据各自效用及资源数量进行博弈, 最终达到最优分配. 因此, 与 CRP 和 SMP 相比, ROAP 策略兼顾了效率和公平. 实验中仅设置了资源供大与求时的情况, 对于资源供不应求的情况, 容易得出每种分配策略各 RJES 均能得到近似收益的结论.

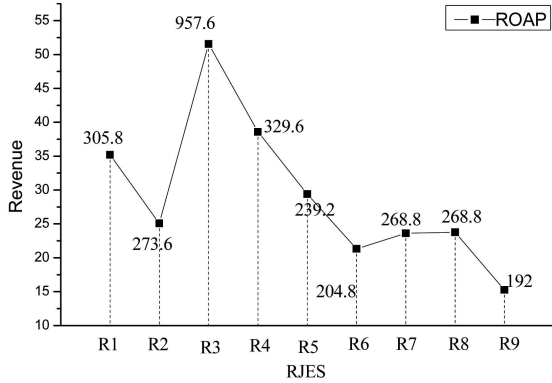


图 3 ROAP 策略下各 RJES 收益及其能力比较

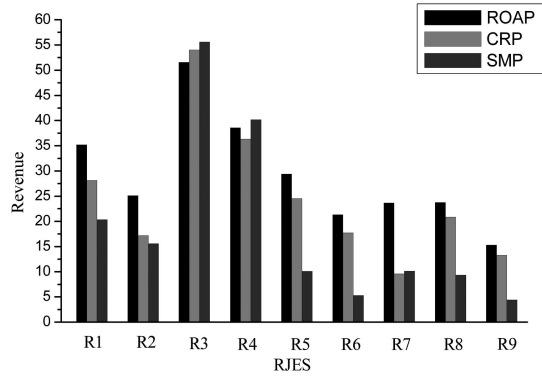


图 4 3 种分配策略产生的收益比较

#### 2) 服务拒绝率比较

本文通过冗余机制来增强任务执行的健壮性, 得出了节点的不同信任度条件下 3 种分配策略的作业拒绝率比较情况. 如图 5 所示, 第 1 组 (1st grp) 信任度为表 1 中指定的信任值, 第 2 组 (2nd grp) 为表 1 中信任值的 0.9 倍, 第 3 组 (3rd grp) 为表 1 中信任值的 0.8 倍. 可以看出, 根据 ROAP 策略进行任务调度时作业拒绝率最低. 并且随着节点信任度的减少, ROAP 策略优势越来越大, 与 CRP 策略相比最高时可降低 69%, 与 SMP 策略相比最高时可降低 66%.

RJES 的信任度阈值对作业拒绝率和 RJES 中节点组合均产生影响. 阈值越大, 作业拒绝率越低, 用户 QoS 保障越好, 但 RJES 中节点数目却越多, 资源利用率越低. 因此, 在用户 QoS 保障和资源利用率之间存在一定的 “trade-off”. 图 6 显示了 RJES 的信任度阈值对作业拒绝率的影响, 随着阈值的增大, 作业拒绝率下降减慢, 阈值为 0.92 时的作业拒绝率与阈值为 0.9 时相比仅下降了 8.4%. 大量实验表明, 信任度阈值在 0.9 左右时用户 QoS 能得到比较满意的保障. 对于如何在资源利用率与用户 QoS 保障之间取得平衡将在下一步工作中进行深入探讨.

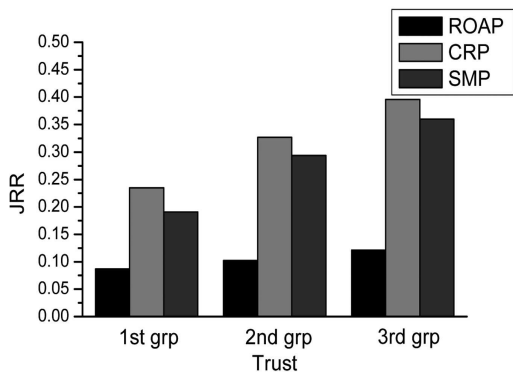


图 5 不同信任条件下作业拒绝率比较

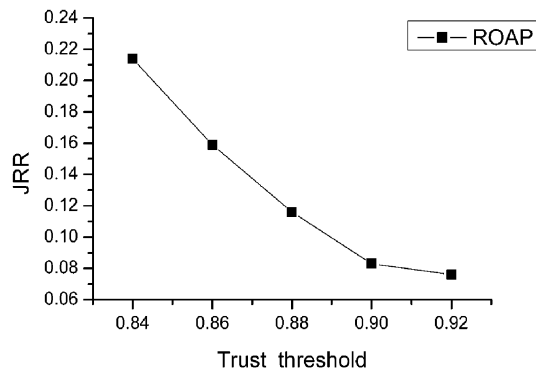


图 6 信任度阈值对作业拒绝率的影响

### 3) 信任评估检验

对网格节点信任行为进行模拟, 假设节点  $S_1$ 、 $S_4$  信任行为变好. 在市场供大于求的情况, 仅根据信任度来选择资源会造成节点  $S_1$ 、 $S_4$  信任度无法上升, 难以客观评估这两个节点的信任行为, 从而使其收益减少, 减弱其加入网络的积极性. 本文的 RMROA 模型通过冗余技术增加资源执行任务的机会, 提高了节点的信任度, 较准确地刻画了一个资源的信任行为. 图 7 显示了节点  $S_1$ 、 $S_4$  的信任变化情况, 随着执行任务历史记录 (Number of records) 的增加, 节点信任度提高并逐渐平稳, 与节点的真实行为相符.

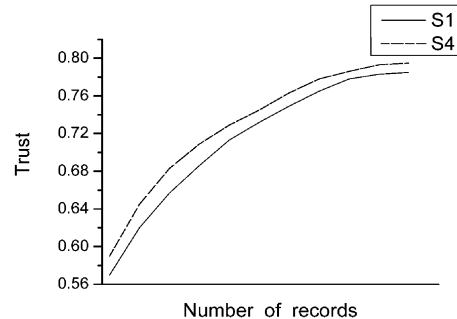


图 7 网格节点信任变化

## 6 结束语

用户 QoS 保障是网格的重要目标之一, 本文研究集中在不可靠的网格环境下增强 QoS 保障的资源分配问题. 提出了健壮性增强的作业执行服务的概念及其确定方法, 基于该概念提出了健壮性增强的离线资源分配模型, 通过将 RJES 中多个网格节点中资源分配给作业冗余执行以增强其健壮性. 基于该模型, 将资源分配描述为一个非合作博弈过程, 求得了帕累托最优分配方案, 并提出了一种健壮性增强的资源分配策略. 实验结果表明本文提出分配模型兼顾了效率和公平, 增强了作业执行的健壮性, 并较好地评估资源的信任行为. 下一步工作将主要集中在: 1) 如何在用户 QoS 保障与资源利用率之间取得平衡; 2) RJES 中各网格节点收益如何分配.

## 参考文献

- [1] Foster I, Kesselman C. The Grid: Blueprint for a New Computing Infrastructure[M]. San Francisco: Morgan Kaufmann Publishes, 2004.
- [2] Buyya R. Economic-based Distributed Resource Management and Scheduling for Grid Computing[M]. Monash: Monash University, 2002.
- [3] Czajkowski K, Foster I, Kesselman C. Resource co-allocation in computational grids[C]//Proc of the 8th IEEE International Symposium on High Performance Distributed Computing, 1999: 219-228.
- [4] Dai Y S, Levitin G, Trivedi K S. Performance and reliability of tree-structured grid services considering data dependence and failure correlation[J]. IEEE Transactions on Computers, 2007, 56(7): 925-936.
- [5] Sodan A C, Doshi C, Barsanti L, et al. Gang scheduling and adaptive resource allocation to mitigate advance reservation impact[C]//Proc of the 6th IEEE International Symposium on Cluster Computing and the Grid, 2006: 5.
- [6] Wu M, Sun X H, Chen Y. QoS oriented resource reservation in shared environments[C]//Proc of the 6th IEEE International Symposium on Cluster Computing and the Grid, 2006: 601-608.
- [7] 胡志刚, 胡周君, 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.  
Hu Z G, Hu Z J, Li L. Performance evaluation based meta task scheduling algorithm[J]. Journal of Xi'an Jiaotong University, 2008, 42(8): 972-976.
- [8] 胡志刚, 胡周君. 计算服务网格中基于服务聚类的元任务调度算法 [J]. 小型微型计算机系统, 2009, 30(2): 248-254.  
Hu Z G, H Z J. Service-clustering-based meta-task scheduling algorithm in computational grids[J]. Journal of Chinese Computer Systems, 2009, 30(2): 248-254.
- [9] Hu Z J, Hu Z G, Liu Z H. Resource availability evaluation in service grid environment[J]. 2007 IEEE Asia-Pacific Services Computing Conference, 2007: 232-238.



- [10] 胡志刚, 胡周君. 一种基于预测的资源匹配算法 [J]. 计算机应用, 2007, 27(10): 2391–2394.  
Hu Z G, Hu Z J. Prediction-base resource matching algorithm[J]. Journal of Computer Applications, 2007, 27(10): 2391–2394.
- [11] Lewis M, et al. Support for extensibility and site autonomy in the legion grid system object model[J]. Journal of Parallel and Distributed Computing, 2003, 63(5): 525–538.
- [12] Foster I, Kesselman C, Lee C, et al. A distributed resource management architecture that supports advance reservation and co-allocation[C]//Proc of the 7th IEEE International Workshop on Quality of Service, 1999: 27–36.
- [13] Waldrich O, Wieder P, Ziegler W. A meta-scheduling service for co-allocating arbitrary types of resources[R]. CoreGRID Technical Report TR-0010, 2005.
- [14] Azougagh D, Yu J L, Kim J S, et al. Resource co-allocation: A complementary technique that enhances performance in grid computing environment[C]//Proc of the 11th International Conference on Parallel and Distributed System, 2005: 36–42.
- [15] Abramson D, Giddy J, Foster I, et al. High performance parametric modeling with nimrod/G: Killer application for the global grid?[C]//Proc of the International Symposium on Parallel and Distributed Processing, 2000: 520–528.
- [16] Ghosh P, Roy N, Das S K, et al. A game theory based pricing strategy for job allocation in mobile grids[C]//Proc of the 18th International Parallel and Distributed Processing Symposium, 2004: 82.
- [17] Kwok Y K, Hwang K, Song S. Selfish grids: Game-theoretic modeling and NAS/PSA benchmark evaluation[J]. IEEE Transactions on Parallel and Distributed Systems, 2007, 18(5): 621–636.
- [18] Rzacca K, Trystram D, Wierzbicki A. Fair game-theoretic resource management in dedicated grids[C]//Proc of the 7th IEEE International Symposium on Cluster Computing and the Grid, 2007: 343–350.
- [19] Tan Z, Gurd J R. Market-based grid resource allocation using a stable continuous double auction[C]//8th IEEE/ACM International Conference on Grid Computing, 2007: 283–290.
- [20] 翁楚良, 陆鑫达. 一种基于市场机制的网格资源调价算法 [J]. 计算机研究与发展, 2004, 41(7): 1151–1156.  
Weng C L, Lu X D. A pricing algorithm for market-based resource management on grid computing systems[J]. Journal of Computer Research and Development, 2004, 41(7): 1151–1156.
- [21] Khan S U, Ahmad I. Non-cooperative, semi-cooperative, and cooperative games-based grid resource allocation[C]//Proc of the 20th International Parallel and Distributed Processing Symposium, 2006: 10.
- [22] 彭冬生, 林闯, 刘卫东. 一种直接评价节点诚信度的分布式信任机制 [J]. 软件学报, 2008, 19(4): 936–945.  
Peng D S, Lin C, Liu W D. A distributed trust mechanism directly evaluating reputation of nodes[J]. Journal of Software, 2008, 19(4): 936–945.
- [23] Cormen T H, Leiserson C E, Rivest R L, et al. Introduction to Algorithms[M]. Second Edition, MIT Press, 2001.
- [24] Tang A, Wang J, Low S H. Counter-intuitive throughput behaviors in networks under end-to-end control[J]. IEEE/ACM Transactions on Networking, 2006, 14(2): 355–368.
- [25] Nash J. Non-cooperative games[J]. The Annals of Math, 1951, 54.
- [26] 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: 1175–1220.
- [27] Cappello F, et al. Grid'5000: A large scale and highly reconfigurable experimental grid testbed[J]. International Journal of High Performance Computing Applications, 2006, 20( 4): 481–494.
- [28] Iosup A, Jan M, Sonmez O, et al. On the dynamic resource availability in grids[C]//8th IEEE/ACM International Conference on Grid Computing, 2007: 26–33.
- [29] Iosup A, Sonmez O, Epema D. The characteristics and performance of groups of jobs in grids[J]. Lecture Notes in Computer Science, 2007, 46(41): 382–393.
- [30] 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.