

文化框架下多群智能优化算法的云作业调度

孙琼琼, 蔡琪

(平顶山教育学院 计算机科学与技术系, 河南 平顶山 467002)

摘要: 作业调度是一种云计算核心技术, 为了获得更优的云计算作业调度方案, 提出一种文化框架下多群智能优化算法的云作业调度方法; 首先构建云作业调度问题的数学模型, 然后借助文化算法模型, 粒子群算法组成信仰空间, 人工鱼群算法组成群体空间, 两者之间并行演化, 相互促进, 对云计算作业调度数学模型进行求解, 最后通过仿真实验测试算法的性能; 结果表明, 本文加快了算法的收敛速度, 获得了更优的云计算作业调度方案, 大幅度缩短少云计算作业完成时间, 具有一定的实用价值。

关键词: 云计算; 作业调度; 人工鱼群算法; 文化算法; 粒子群算

Job Scheduling in Cloud Computing Based on Multi-swarm Optimization Algorithm With Cultural Algorithm

Sun Qiongqiong, Cai Qi

(Department of Computer Science, Pingdingshan Institute of Education, Pingdingshan 467002, China)

Abstract: Job scheduling is the core technology of cloud computing, in order to obtain good scheduling results, a new job scheduling method in cloud computing based on multi-swarm optimization algorithm with cultural algorithm is proposed in this paper. Firstly, The mathematical model is constructed for cloud scheduling problem, and then with the help of cultural algorithm model, particle swarm algorithm consists of the belief space while artificial fish swarm algorithm consists of group space, their parallel evolution and promote each other to solve the job scheduling mathematical model, finally, the simulation experiment is carried out to test the performance of the algorithm. The results show that, the proposed algorithm has fastened convergence speed and can obtain a better job scheduling scheme of cloud computing, and greatly shorten the work time, so it has a certain practical value.

Keywords: job scheduling; cloud computing; artificial fish swarm algorithm; cultural algorithm; particle swarm algorithm

0 引言

由于每一个用户的需求不同, 导致作业类型的多样性, 云计算系统如何为各类用户提交的作业分配合理的资源, 进行有效调度, 为用户提供满意的服务, 已经成为云计算研究中的一个重要课题^[1-3]。

针对作业调度问题, 国内外学者和研究人员进行了广泛而深入的研究, 提出一些云计算作业调度算法^[4]。传统云计算作业调度算法主要包括: 公平调度算法、先入先出队列 (first input first output, FIFO) 算法等^[5-6], 这些算法实现简单, 适合于小规模云计算作业问题求解, 但是存在资源浪费、响应时间长等缺陷, 不能满足现代大规模云计算作业调度的需求^[7]。大量研究表明, 云计算作业调度是一个 NP 难问题, 传统算法难以对其进行有效求解^[8]。近年来, 随着群智能算法的不断发展和成熟, 一些学者将其引入到云计算作业调度求解过程中, 出现基于蚁群算法、蛙跳算法、人工鱼群算法、粒子群算法、遗传算法等云作业调度方法, 它们将云计算作业调度看作是一个多条件约束下, 以作业总执行时间最短作为优化目标, 然后进行求解, 找到用户满意度最高的云作业调度方案,

研究表明, 相对于传统算法, 群智能算法加快了求解速度, 大幅度减少了任务的响应时间, 群智能算法成为云作业调度问题的主要研究方向^[9-11]。然而在实际应用中, 单一算法群智能存在一些不足, 如易陷入局部最优, 难以找到全局最优云作业调度方案等^[12]。

为了获得更优的作业调度方案, 提出一种文化框架下多群智能优化算法的云作业调度方法。结果表明, 文化框架下多群智能优化算法可以获得了更优的云作业调度方案, 缩短了任务完成时间, 大幅度提高了云计算系统的整体性能和运行效率。

1 云计算作业调度问题描述

1.1 云计算系统体系构架

云计算是通过计算机网络将存储和计算处理程序分配到分布式计算机中, 然后用户可以资源切换到自己需要的应用上。云计算系统体系构架如图 1 所示^[13]。

1.2 云计算作业调度的数学模型

Hadoop 平台是一种分布式计算开源框架, 其核心思想来源于 Google 的 Map/Reduce 模型, 每次计算请求被称为作业, 作业被划分成多个任务执行, 任务分为 Map 任务和 Reduce 任务。图 1 显示了 Map-Reduce 模型的结构框架^[14]。

设用户提交了 $NumT$ 个作业, 第 t 个作业被划分为 M 个 Map 子作业和 R 个 Reduce 子作业, 作业 t 的运行时间 $TaskT$, 定义为:

收稿日期: 2014-06-15; 修回日期: 2014-07-14;

基金项目: 河南省科技计划重点项目(102102210416)

作者简介: 孙琼琼(1980-), 女, 河南郟县人, 讲师, 硕士, 主要从事云计算及计算机网络技术方向的研究。

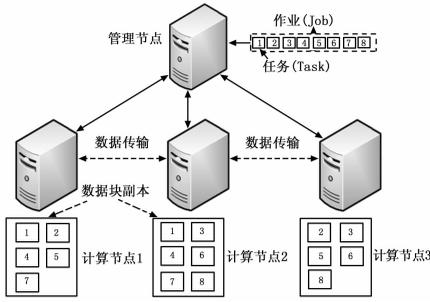


图 1 云计算系统体系架构

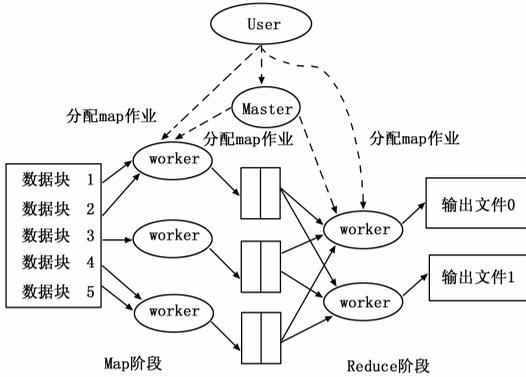


图 2 Map-Reduce 模型

$$TaskT_i = \max(\sum_{j=1}^{Nom} WM(i, j) + \sum_{x=1}^{Nor} WR(k, x)) \quad (1)$$

式中, $WM(i, j)$ 表示第 j 个任务所用的时间; Nom 和 Nor 表示队列中的位置。

作业进行调度的目标为: 总完成时间 T_1 和平均完成时间 T_2 最小, 即

$$T_1 = \min(\sum_{i=1}^{NumT} TaskT_i) \quad (2)$$

$$T_2 = \min(\frac{1}{NumT} \sum_{i=1}^{NumT} TaskT_i) \quad (3)$$

从式 (2)、(3) 可知, 云计算作业调度问题变成了最小值优化问题, 作业调度的数学模型为:

$$T = \min(\sum_{i=1}^{NumT} TaskT_i) + \min(\frac{1}{NumT} \sum_{i=1}^{NumT} TaskT_i) \quad (4)$$

本文采用文化框架下多群智能优化算法对式 (4) 的云计算作业调度问题进行求解。

2 文化算法框架的多群智能算法

由于粒子群算法具有收敛速度快, 局部寻优能力强的优点, 而且人工鱼群法具有强大的全局搜索能力, 但是它们均存在息的不足, 为了克服单一算法的不足, 文化算法通过将它们分别组成信仰空间和群体空间, 组成一新的多群智能算法。

2.1 人工鱼群算法

在人工鱼群算法, 人工鱼作为真实鱼的一个虚拟实体, 问题解的空间是人工鱼所处的环境及人工鱼所处的状态, 人工鱼下一时刻的行为取决于目前自身和目前环境的状态, 通过自身行为活动影响同伴的活动, 进而适应环境。人工鱼群通过模拟鱼类的 4 种行为: 觅食行为、群聚行为、追尾行为和随机行

为, 这些行为在不同的条件下可互相转换, 通过人工鱼算子及启发函数来引导人工鱼寻优, 选择当前最优的行为执行, 使鱼群聚集于食物浓度最高的位置, 此也即所求问题的解。相关定义如下。

定义 1: 人工鱼个体向量 $\mathbf{X} = (a_1, a_2, \dots, a_n)$, 其中 a_i 表示要寻优的变量。

定义 2: 人工鱼当前所处位置的食物浓度 $F(X_i)$ (目标函数数值)。

定义 3: 为配合模拟人工鱼的行为, 定义 3 个控制参数: 人工鱼的视野 (*visual*)、鱼群的拥挤度因子 (δ) 及觅食尝试次数 (*try_number*)。

定义 4: 人工鱼 X_1 和 X_2 的距离定义为 X_1 和 X_2 中对应不相等分量的个数, 记为 $Distance(X_1, X_2)$ 。

定义 5: 人工鱼 \mathbf{X} 的 K 邻域定义为 $N(\mathbf{X}, K) = \{X' \mid Distance(\mathbf{X}, X') < k, X' \in D\}$, 其中 D 表示变量空间, 若 $X' \in N(\mathbf{X}, K)$, 则 X' 为 \mathbf{X} 的 K 邻域。

定义 6: 人工鱼邻域中心定义为 X 邻域内人工鱼对应分量均值取整 $X_c = \frac{1}{nf} \sum_{i=1}^{nf} X_i$, 其中 nf 表示 X 的 K 邻域内人工鱼个数。

2.2 粒子群算法

粒子群优化 (PSO) 算法模拟鸟类的群体飞行觅食行为, 在 D 维搜索空间中, 第 i 个粒子 ($i = 1, 2, \dots, m$) 的位置为 $Z_i = (z_{i1}, z_{i2}, \dots, z_{iD})$, $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ 为粒子 i 的位置移动距离, $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ 表示第 i 个粒子“飞行”历史中最优位置, $P_g = (p_{g1}, p_{g2}, \dots, p_{gD})$ 表示种群历史最优位置, 粒子根据以下公式更新速度和位置:

$$v_{id}(t+1) = \omega \times v_{id}(t) + c_1 \times rand() \times (p_{id}(t) - z_{id}(t)) + c_2 \times rand() \times (p_{gd}(t) - z_{id}(t)) \quad (5)$$

$$z_{id}(t+1) = z_{id}(t) + v_{id}(t+1) \quad (6)$$

式中, t 表示迭代次数; c_1, c_2 为学习因子; $rand()$ 是 0~1 之间的随机数; ω 是惯性权重。

2.3 文化算法框架的人工鱼群算法和粒子群算法融合

文化算法可以将其它群体算法融合于文化的框架内, 克服各自的不足, 因此本文的群体空间采用人工鱼群算法, 信仰空间采用粒子群算法, 得到一种文化算法框架下的多群智能算法。

2.3.1 群体空间的设计

演化操作采用人工鱼群算法, 对其进行相应改进: 省略算法中拥挤度因子的计算, 在觅食行为中, 让人工鱼直接移动到搜索到的感知范围内的较好位置。在算法的随机行为中添加一个随机数 r , 每迭代一次 r 在 $[0, 1]$ 之间随机产生, 当 r 大于 $pselet$ 时, 采取向当前全局最优鱼位置前进的行为; 当 r 小于或等于 $pselet$ 时, 采取随机行为。使 *try_number* 随着迭代线性递增, 即

$$try_number = fix(a * k + b) \quad (7)$$

式中, k 为当前迭代次数, a, b 为常数, fix 为取整函数。

2.3.2 信仰空间的设计

演化操作采用粒子群算法, 惯性权重采取线性递减策略,

种群规模取群体空间的 20%, 在信仰空间增加变异操作, 变异率为:

$$\begin{cases} p_m = p_1 - \frac{p_2 * (f(i) - f_{mean})}{(f_{max} - f_{min})}, f(i) < mean \\ p_m = p_1, f(i) \geq mean \end{cases} \quad (8)$$

式中, p_1, p_2 为常数, $f(i)$ 为当前粒子的适应值, $f_{min}, f_{mean}, f_{max}$ 分别为当前种群最差、平均以及最佳粒子的适应值。

变异粒子的位置为:

$$bypos = pos + p_m * \exp(f(i)) * rand \quad (9)$$

式中, pos 为初始粒子的位置, \exp 为指数函数, $rand$ 为 0 与 1 之间之间的随机数。

2.3.3 接受操作

在群体空间的演化过程中, 每运行 $acceptStep$ 代时, 用当前 $bestfish$ 替换信仰空间中群体的最差个体, 本文 $acceptStep = 10$ 。

2.3.4 影响操作

群体空间运行每间隔 $affectStep$ 代时将信仰空间中适应值较好的一部分个体替换群体空间中适应值较劣的同样数目个体。 $affectStep$ 的计算公式为:

$$affectStep = N_1 - N_2 * (i - 1) \quad (14)$$

式中, N_1, N_2 为常数, i 为由 1 到 $fix(N_1/N_2 + 1) + 1$ 逐渐增加的整数。

这样在群体空间演化的初期, 信仰空间对其影响较小, 保证其快速演化, 在群体空间演化的后期, 可以更多地接受信仰空间的引导, 扩大搜索范围, 提高全局寻优能力。

2.4 多群智能算法的云作业调度求解流程

粒子群收敛速度快, 局部寻优能力强, 但易出现早熟现象, 人工鱼群算法具备强大的全局搜索能力, 但存在搜索速度慢, 收敛精度差的缺点, 本文通过文化演化算法将两者组合在一起, 应用于云作业调度问题求解, 具体流程如图 3 所示。

3 仿真实验

3.1 仿真环境

为了测试本文算法的性能, 实验环境包括 10 台虚拟机, 其中一台为服务器, 其余的为客户端, 每台虚拟机的 CPU 为 4 核 2.8 GHz, 8 G 内存, 1 TB 的硬盘, 采用 windows 2000 操作系统, 选择公平调度算法、人工鱼群算法、粒子群算法进行对比实验。

3.2 结果与分析

3.2.1 不同作业数量下的性能对比

当云计算作业数量发生变化的情况下, 公平调度算法、人工鱼群算法、粒子群算法以及本文算法的作业完成时间变化曲线如图 4 所示。从图 4 可以清楚看出, 随着作业数的不断增加, 所有算法的作业完成时间相应上升, 其中比公平调度算法的完成时间最长, 而且其它算法完成时间相对较短, 这主要是由于其它算法均是群智能算法, 具有良好的全局搜索能力, 可以找到更优的云作业调度方案。

同时本文算法的作业完成时间均小于对比算法, 这主要是由于本文算法通过文化框架对人工鱼群算法和粒子群算法进行

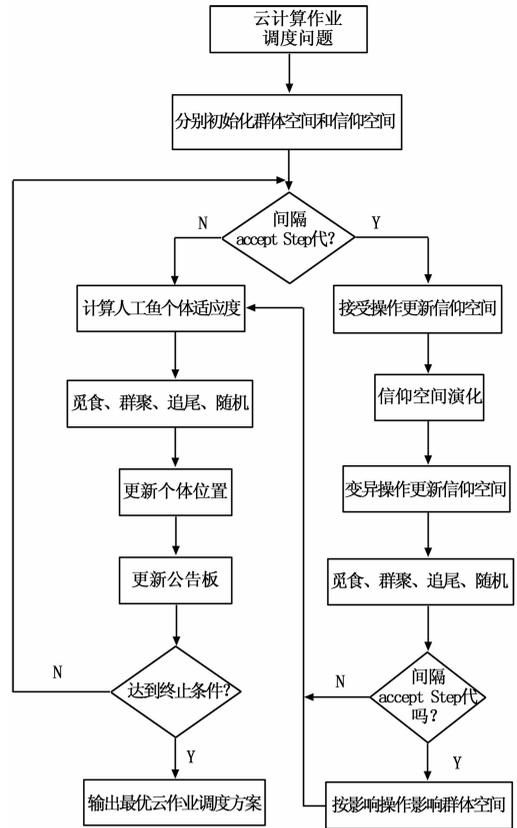
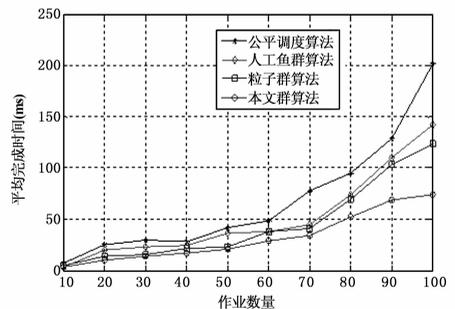
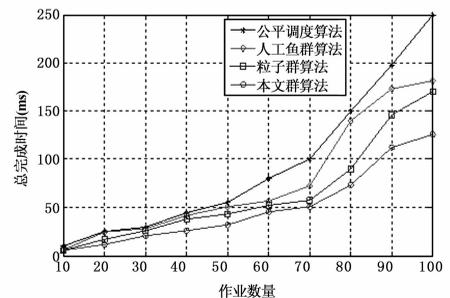


图 3 多群智能算法的作业调度流程

了集成, 较好的克服各自的缺陷, 可以快速找到最优的云作业调度方案, 获得更加理想的调度效果。



(a) 平均完成时间变化曲线



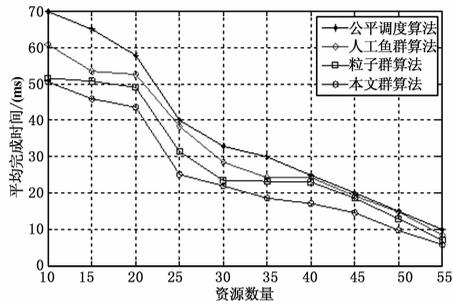
(b) 总完成时间变化曲线

图 4 不同作业数量下的算法性能对比

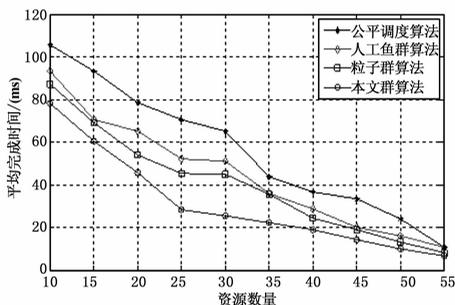
3.2.2 不同资源数量下的性能对比。

当云计算资源数量发生变化的情况下, 公平调度算法、人

工鱼群算法、粒子群算法以及本文算法的作业总完成时间和平均完成时间变化曲线如图 5 所示。从图 5 可知，随着云计算资源数量的增加，作业总完成时间和平均完成时间都逐渐减少，这主要是由于作业竞争资源机会减少，用户作业冲突概率降低，在相同条件下，本文算法获得比了对比算法更短的作业完成时间，对比结果表明，本文算法集成了粒子群算法和人工鱼群算法的优点，加快算法的收敛速度和精度，可以充分利用云计算资源，提高了云计算作业的调度效率，可以较好地满足用户的需求。



(a) 平均完成时间变化曲线



(b) 总完成时间变化曲线

图 5 不同云计算资源数量下的算法性能对比

3.2.3 用户的满意度对比

公平调度算法、人工鱼群算法、粒子群算法以及本文算法的用户满意度如图 6 所示。从图 6 可知，本文算法的平均用户满意度均为 1，这表明表示用户获得了期望的资源分配方式，其作业得到的较好的完成，而其它算法的满意度比较低，结果再一次证明了本文算法的优越性。

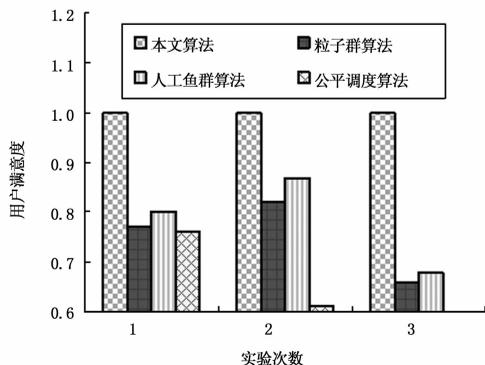


图 6 不同算法的用户满意度对比

4 结束语

针对单群智能算法存在收敛速度慢，易获得局部最优方案的问题，提出了一种文化框架下多群智能算法的云计算作业调度方法，该方法借助文化算法模型，由粒子群算法组成信仰空间，人工鱼群算法组成群体空间，它们并行演化，相互促进，共同完成云计算作业调度问题的求解，最后通过仿真实验测试算法的性能。结果表明，本文算法不仅提高云计算资源的利用率，提高大幅度减少了云作业的执行时间，获得更加合理的云计算作业调度方案。

参考文献:

- [1] Armbrust M, Fox A, Griffith R, et al. A view of cloud computing [J]. Communications of the ACM, 2010, 53 (4): 50-58.
- [2] Mell P, Grance T. The NIST definition of cloud computing [J]. Communications of the ACM, 2010, 53 (6): 50.
- [3] Buyya R, Yeo C S, Venugopal S, et al. Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility [J]. Future Generation Computer Systems, 2009, 25 (6): 599-616.
- [4] Dean J, Ghemawat S. Map-Reduce: Simplified data processing on large clusters [J]. Communications of the ACM, 2012, 51 (1): 107-113.
- [5] Braun T D, Siegel H J, Beck N, et al. 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.
- [6] Zaharia M, Borthakur D, Sarma J S. Job scheduling for multi user Map-Reduce clusters [A]. Proceedings of the 5th European Conference. Washington: IEEE Computer Society [C]. 2012, 10: 145-161.
- [7] Zahafia M, Konwinski A, Joseph A. Improving Map-Reduce performance in heterogeneous environments [A]. Proceedings of the 8th UNIX Operating Systems Design and Implementation [C]. 2011, 8: 29-42.
- [8] In J, Luo J, Song A, et al. BAR: an efficient data locality driven task scheduling algorithm for cloud computing [A]. Proceedings of the CCGRID [C]. 2011, 11: 295-304.
- [9] 陈全, 邓倩妮. 异构环境下自适应的 Map-Reduce 调度 [J]. 计算机工程与科学, 2009, 31 (1): 101-105.
- [10] 余正祥. 基于学习方式对 Map-Reduce 作业调度的改进研究 [J]. 计算机科学, 2012, 39 (6A): 220-222.
- [11] 华夏渝, 郑骏, 胡文心. 基于云计算环境的蚁群优化计算资源分配算法 [J]. 华东师范大学学报 (自然科学版), 2010, 10 (1): 127-134.
- [12] 高丽丽, 刘弘, 李同喜. 基于文化粒子群算法的约束优化问题求解 [J]. 计算机工程, 2008, 5 (3): 179-181.
- [13] 李建锋, 彭舰. 云计算环境下基于改进遗传算法的任务调度算法 [J]. 计算机应用, 2011, 31 (1): 184-186.
- [14] 梁静, 许波, 葛宇. 基于改进蛙跳策略的 Map-Reduce 作业调度算法 [J]. 计算机应用研究, 2013, 30 (7): 1999-2002.