

基于改进量子遗传算法的云计算资源调度

刘卫宁^{1,2}, 靳洪兵^{1,2*}, 刘波^{1,2}

(1. 重庆大学 计算机学院, 重庆 400044; 2. 信息物理社会可信服务计算教育部重点实验室(重庆大学), 重庆 400044)

(* 通信作者电子邮箱 jasonjin2009@163.com)

摘要: 针对云计算环境下资源的高效调度问题, 当前研究较少关注云服务提供商的服务成本, 为此, 以云服务提供商降低最小服务成本为目的, 提出了改进量子遗传算法的云资源调度算法。由于采用二进制量子位表示的染色体无法描述资源调度矩阵, 该算法将量子位的二进制编码转换为实数编码, 并使用旋转策略和变异算子保证算法的收敛性。通过仿真实验平台将此算法与遗传算法和粒子群算法进行比较分析, 在种群迭代次数为 100 的情况下, 分别取种群数为 1 和 10, 实验结果表明该算法能取得更小的最小服务成本。

关键词: 云计算; 量子遗传算法; 资源调度; 最小成本; 实数编码

中图分类号: TP18; TP393.027 **文献标志码:** A

Cloud computing resource scheduling based on improved quantum genetic algorithm

LIU Weining^{1,2}, JIN Hongbing^{1,2*}, LIU Bo^{1,2}

(1. College of Computer Science, Chongqing University, Chongqing 400044, China;

2. Key Laboratory of Dependable Service Computing in Cyber Physical Society, Ministry of Education (Chongqing University), Chongqing 400044, China)

Abstract: Focusing on the problem of high efficiency resource scheduling in cloud computing environment, since current research has been less concerned about the cost of the services of the cloud service provider, an improved Quantum Genetic Algorithm (QGA) was proposed to reduce the minimum service cost of cloud service provider. This algorithm converted quantum-bits encoded by binary number to real-coded quantum-bits as chromosome represented by binary-coded quantum-bits cannot describe the resource scheduling matrix, and used rotation strategy and mutation operator to guarantee the convergence of the algorithm. Comparative experiments were conducted among the improved QGA, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) through simulation platform, the populations number is 1 and 100 with 100 iteration times. The experimental results show that the improved QGA can obtain smaller minimum service cost.

Key words: cloud computing; Quantum Genetic Algorithm (QGA); resource scheduling; minimum cost; real encoding

0 引言

云计算是一种利用互联网实现随时随地、按需、便捷的访问共享资源池(如计算设施、存储设备、应用程序等)的计算模式^[1]。云计算按服务层级可分为三种服务架构:1)基础设施及服务(Infrastructure as a Service, IaaS);2)平台及服务(Platform as a Service, PaaS);3)软件及服务(Software as a Service, SaaS)。

本文重点关注 IaaS 层服务架构, 由于基础设施的异构性, 合理的资源调度方式成为影响云服务质量的关键。为了最大限度地满足用户的服务需求, 又产生资源浪费, 云计算资源调度问题就成为学界关注的焦点。

针对资源调度问题的求解, 主要有基于经典优化问题的求解技术, 即将调度问题归结为某种已知的数学问题, 在此基础上构造模型, 并采用该数学问题经典的求解方法。但由于资源调度问题是一个 NP-hard 问题, 在解空间规模迅速增加的情况下, 该方法因其较高的计算时间复杂性会导致其性能的严重局限。因而, 智能算法越来越多地应用到资源调度问题的求解中, 例如, 遗传算法(Genetic Algorithm, GA)^[2-4]。

而云计算降低了资源准入的门槛, 汇聚了更多的资源, 形成了海量的候选资源池, 极大地增加了云计算环境下资源调度的问题规模, 同时也对调度问题的求解效率提出了严峻的挑战。量子遗传算法(Quantum Genetic Algorithm, QGA)因其具有按概率优化和广阔的解空间搜索特性而被广泛应用^[5-6]。但尚未用于对云计算资源调度问题的求解, 而且, 既有的量子遗传算法采用二进制编码, 对大规模云计算资源调度问题的求解尚存在局限。

鉴于此, 本文提出基于实数编码的改进量子遗传算法, 利用量子算法的按概率优化特性, 使用旋转门更新, 并增加变异算子, 使得在量子编码的广阔空间中避免获得局部最优解。

1 云资源调度模型

1.1 调度模型

云计算环境下的资源调度模型可以用 n 个用户和 m 个资源组成的四元组来描述^[7]:

$$M = (U, V, F, \theta)$$

其中: U 为 n 个用户所组成的用户集合; V 为 m 个虚拟机所组成的虚拟机资源集合; F 为云资源调度优化的目标函数; θ 则

收稿日期: 2013-02-04; 修回日期: 2013-03-17。

基金项目: 国家自然科学基金资助项目(61203135); 国家科技支撑计划项目(2012BAH19F01)。

作者简介: 刘卫宁(1965-), 女, 重庆人, 教授, 博士生导师, 主要研究方向: 智能计算与服务、物流与供应链管理、网络与分布式计算; 靳洪兵(1988-), 男, 重庆人, 硕士研究生, 主要研究方向: 云计算、智能计算; 刘波(1981-), 男, 四川广安人, 博士研究生, 主要研究方向: 先进制造、供应链管理优化。

表示调度优化算法,多采用智能算法。本文只考虑一个用户所执行任务的情况,多个用户可以循环执行。调度模型的具体特征描述如下:

1) 虚拟机资源集合 $V = \{v_1, v_2, \dots, v_m\}$, 每个虚拟机按照内核数、内存大小和磁盘空间大小进行细化。 $v_i = \{\lambda_i, \mu_i, \varphi_i\} (i = 1, 2, \dots, m)$, 其中: i 表示虚拟机编号; λ_i, μ_i 和 φ_i 分别表示虚拟机内核数、内存大小和磁盘空间大小。

2) 元任务集 $T = \{t_1, t_2, \dots, t_n\}$, 表示一个用户的 n 个相互独立的元任务。

3) 任务时间矩阵 $S_{n \times m} = (s_{ij}) (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$, s_{ij} 是任务 i 在资源 j 上的执行时间。任务时间矩阵根据分配的虚拟机性能大小取值。

4) 资源调度矩阵 $E_n = (e_i) (i = 1, 2, \dots, n)$, 表示资源 e_i 执行任务 i 。若 $E_n = [3 \ 2 \ 2 \ 4 \ 1]$, 表示3号虚拟机执行任务1, 2号虚拟机执行任务2, ..., 1号虚拟机执行任务5。

5) 资源使用矩阵 $X_{n \times m} = (X_{je_i})$, 其中: $j = 1, 2, \dots, n, e_i \in E_n, X_{je_i} = 1$, 其余为0。 $X_{n \times m}$ 是根据资源调度矩阵产生的一个资源使用矩阵, X_{je_i} 表示第 e_i 号虚拟机被任务 j 使用。

本文针对资源调度模型作出如下假设:

假设1 虚拟机的性能能够满足任一任务的要求;

假设2 一个任务只能分配给一个虚拟机^[8-9];

假设3 所有的任务可以完全分配^[8-9]。

由任务时间矩阵和资源使用矩阵可以得出虚拟机 v_j 上任务完成所用时间:

$$T_j = \max_{1 \leq i \leq n} \{s_{ij} \cdot x_{je_i}\}; j = 1, 2, \dots, m, i = 1, 2, \dots, n \quad (1)$$

从而, m 个虚拟机完成任务所使用总时间为:

$$makespan = \sum_{j=1}^m \max_{1 \leq i \leq n} \{s_{ij} \cdot x_{je_i}\}; j = 1, 2, \dots, m, i = 1, 2, \dots, n \quad (2)$$

1.2 云服务成本

在云计算环境中,服务提供商不仅关心用户提交任务的完成时间,还关心提供云服务的成本。云计算的 Pay-as-you-go 服务模式要求提供商对用户的服务进行量化收费,本文将云资源的虚拟机作为调度单位,用户任务需求按照虚拟机实例来进行分配调度。针对虚拟机的不同配置,用户任务在虚拟机上的执行时间也会相应不同,因此良好的调度算法可以减少服务提供商的云服务成本,从而增加服务提供商的收益。本文借鉴 Google Compute Engine 的虚拟机定价模式^[10], 制定出云服务提供商虚拟机的单位时间成本 p_j :

$$p_j = p_{i1} \times \alpha_i + p_{i2} \times \beta_i + p_{i3} \times \gamma_i \quad (3)$$

其中: p_{i1}, p_{i2} 和 p_{i3} 分别为单位时间内虚拟机单位内核、内存以及磁盘空间的成本^[9]。那么每个虚拟机的单位时间成本为

$$P = \{p_1, p_2, \dots, p_m\} \quad (4)$$

针对用户的某一任务,云服务提供商所需承担的虚拟机 v_j 成本为

$$P_j = (p_{i1} \times \alpha_i + p_{i2} \times \beta_i + p_{i3} \times \gamma_i) \times \max_{1 \leq i \leq n} \{s_{ij} \cdot x_{je_i}\} = p_j \times \max_{1 \leq i \leq n} \{s_{ij} \cdot x_{je_i}\} \quad (5)$$

式(5)由式(3)与式(2)的乘积得到。那么,针对一个用户提交的任务,云服务商执行总成本为

$$P = \sum_{j=1}^m p_j \times \max_{1 \leq i \leq n} \{s_{ij} \cdot x_{je_i}\} \quad (6)$$

将所有虚拟机的成本和与式(2)相乘可得云服务执行总成本 P 。由此,确定本文的目标函数为: $\min P$ 。

2 量子遗传算法

2.1 算法原理

量子遗传算法是基于量子计算原理的概率优化方法^[11]。它将量子比特的概率幅表示方式应用于染色体的编码,使得一个染色体具有多个模态的叠加,从而比传统的算法更具有并行性^[8]。

在量子遗传算法中,染色体的编码不是用确定的值表示,而是用量子比特来表示。一个量子位可表示成

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

其中 α 和 β 为复数,表示相应状态的概率幅,且 $|\alpha|^2 + |\beta|^2 = 1$ 。量子比特上的运算必须保持该范数, $|\alpha|^2$ 表示 $|0\rangle$ 的概率, $|\beta|^2$ 表示 $|1\rangle$ 的概率。一个量子位可能处于 $|0\rangle$ 或 $|1\rangle$, 也可能处于 $|0\rangle$ 和 $|1\rangle$ 的叠加态。所以,一个量子位比经典比特位包含了更多的信息。

因此,一个长度为 n 的量子染色体 Q 为:

$$Q = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_1 & \beta_2 & \dots & \beta_n \end{bmatrix}$$

其中 $|\alpha_i|^2 + |\beta_i|^2 = 1 (i = 1, 2, \dots, n)$ 。每个量子比特经过测量后可能得到的状态为0或1,因而染色体 Q 能表示 2^n 个不同的经典信息,所以量子遗传算法可以搜索到比经典遗传算法更为广阔的空间。

本文中量子遗传算法的演化规则是通过量子旋转门来实现的,量子旋转门定义^[12]如下:

$$U(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

$$\begin{pmatrix} \alpha_i^{t+1} \\ \beta_i^{t+1} \end{pmatrix} = U(\theta) \begin{pmatrix} \alpha_i^t \\ \beta_i^t \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \alpha_i^t \\ \beta_i^t \end{pmatrix}$$

其中: $[\alpha_i^t \ \beta_i^t]^T$ 为第 t 代量子染色体第 i 个量子位, $[\alpha_i^{t+1} \ \beta_i^{t+1}]^T$ 为变异后第 $t+1$ 代量子染色体第 i 个量子位。

2.2 算法改进

为了适应本文的云计算资源调度模型,对量子遗传算法做了相应的改进:使用转化函数将二进制编码转化为实数编码^[13],对量子种群应用变异算子。图1为改进量子遗传算法流程。

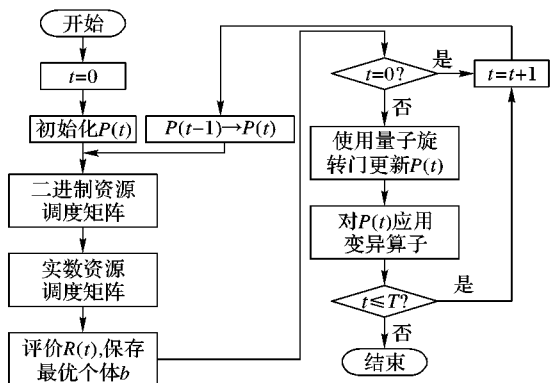


图1 改进量子遗传算法流程

初始种群 $P(t) = \{p_1^t, p_2^t, \dots, p_n^t\}$, 其中 p_j^t 为第 t 代种群中的第 j 个个体。

$$p_j^t = \begin{bmatrix} \alpha_1^t & \alpha_2^t & \dots & \alpha_n^t \\ \beta_1^t & \beta_2^t & \dots & \beta_n^t \end{bmatrix} \quad (7)$$

其中 $j = 1, 2, \dots, n$, n 为量子位数目, 即量子染色体的长度。

$R(t), B(t)$ 表述为: $R(t) = \{\gamma_1^t, \gamma_2^t, \dots, \gamma_n^t\}$, $B(t) = \{b_1^t, b_2^t, \dots, b_n^t\}$ 。其中:

$$\gamma_i^t = \begin{bmatrix} \lceil m \cdot \alpha_1^t \rceil & \lceil m \cdot \alpha_2^t \rceil & \dots & \lceil m \cdot \alpha_n^t \rceil \\ \lceil m \cdot \beta_1^t \rceil & \lceil m \cdot \beta_2^t \rceil & \dots & \lceil m \cdot \beta_n^t \rceil \end{bmatrix} \quad (8)$$

其中: $i = 1, 2, \dots, n$; m 为资源数; b_k^t 是长度为 n 的二进制串, $k = 1, 2, \dots, n$ 。式(8)的产生过程为: 对于式(7), 若 $|\alpha_i^t|^2 > \delta$, 二进制串的第 i 位取 1; 否则, 取 0。 δ 为 0 ~ 1 的一个随机数。

种群数根据情况设置: 量子染色体的长度 n 表示用户提交的作业数, m 为资源数, 即虚拟机个数; $R(t)$ 对应的是资源模型中的资源调度矩阵, γ_i^t 与 b_k^t 一一对应; T 为种群迭代次数, 个体适应度函数为 $(\min P)^{-1}$ 。

除了旋转策略外, 本文还加入了变异算子, 变异算子有变异位数 pn 和变异概率 pm 两参数约束^[14-15]。根据经验值变异概率一般取 0.04, 而变异位数则应根据具体的种群数和编码位数来确定。

3 实验分析

本文采用 CloudSim 作为实验模拟平台, CloudSim 继承了 GridSim 的编程模型, 支持云计算的基础设施的建设与仿真, 本文在 CloudSim 的 DatacenterBroker 中实现了改进的量子遗传算法。

实验分为种群数取 1 和 10, 迭代次数取 100 的情况下, 将本文的改进量子遗传算法 (IQGA) 分别与遗传算法和粒子群算法 (Particle Swarm Optimization, PSO) 进行比较。

从图 2 中可以看出, 当种群数取 1, 迭代次数小于 20 时, IQGA 得到的最小成本比 GA、PSO 取得的最小成本稍大; 但是随着迭代次数的增加, IQGA 取得的最小成本明显比 GA 和 PSO 小。

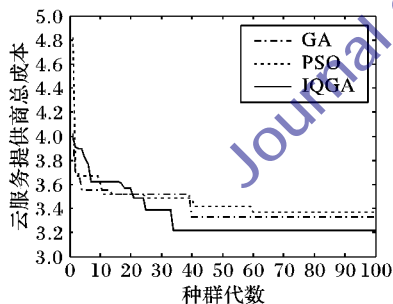


图2 种群数为1时各算法的性能对比(迭代100次)

从图3可以看出, 一开始 IQGA 与 GA 和 PSO 取得的最小成本相差不多, 但当迭代次数大于 50 后, IQGA 的效率就明显比 PSO 和 GA 高, 取得的最小成本也明显更小。

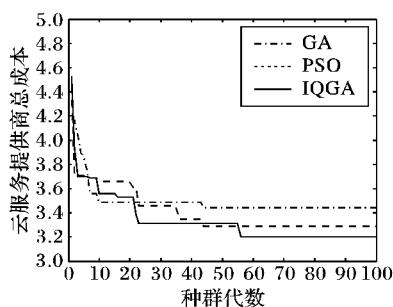


图3 种群数为10时各算法的性能对比(迭代100次)

实验结果表明, 在相同条件下求解云资源调度问题,

IQGA 比 GA 和 PSO 具有更高的效率, 能显著降低云服务商的支付成本。

4 结语

本文对云计算资源调度问题进行了研究, 提出了基于改进量子遗传算法 (IQGA) 的云计算资源调度, 该算法通过改进量子遗传算法的二进制编码使其适用于实数编码。CloudSim 仿真模拟实验结果表明, 在相同条件下求解云资源调度问题, IQGA 比 GA 和 PSO 具有更高的效率, 能显著降低云服务商的支付成本。下一步工作将放宽调度模型的假设条件, 优化算法, 进一步提高求解效率和准确性。

参考文献:

- [1] MELL P, GRANCE T. The NIST definition of cloud computing [EB/OL]. [2012-12-08]. <http://www.nist.gov/itl/cloud/>.
- [2] 李建峰, 彭舰. 云计算环境下基于改进遗传的算法的任务调度算法[J]. 计算机应用, 2011, 31(1): 184-186.
- [3] 齐金平, 查显峰. 多任务多资源优化调度的病毒遗传算法[J]. 计算机应用, 2011, 31(7): 1773-1775.
- [4] 徐晓勇, 潘郁, 凌晨. 云计算环境下资源的节能调度[J]. 计算机应用, 2012, 32(7): 1913-1915.
- [5] LAYEB A, MESHOUL S, BATOCHE M. Multiple sequence alignment by quantum genetic algorithm [C]// IPDPS 2006: Proceedings of the 20th International Parallel and Distributed Processing Symposium. Piscataway: IEEE Computer Society, 2006.
- [6] LI B B, WANG L. A hybrid quantum-inspired genetic algorithm for multiobjective flow shop scheduling [J]. IEEE Transactions on Systems Man and Cybernetics, Part B: Cybernetics, 2007, 37(3): 576-591.
- [7] 孙大为, 常桂然, 李风云, 等. 一种基于免疫克隆的偏好多维 QoS 云资源调度优化算法[J]. 电子学报, 2011, 39(8): 1824-1831.
- [8] ZHENG Z N, WANG R, ZHONG H, et al. An approach for cloud resource scheduling based on parallel genetic algorithm [C]// ICCRD2011: Proceedings of the 3rd International Conference on Computer Research and Development. Washington, DC: IEEE Computer Society, 2011: 444-447.
- [9] YANG Z, YIN C Q, LIU Y. A cost-based resource scheduling paradigm in cloud computing [C]// PDCAT 2011: Proceedings of the 12th International Conference on Parallel and Distributed Computing, Applications and Technologies. Washington, DC: IEEE Computer Society, 2011: 417-422.
- [10] Google compute engine pricing [EB/OL]. [2012-12-20]. <http://cloud.google.com/pricing/compute-engine.html>.
- [11] NARAYANAN A, MOORE M. Quantum-inspired genetic algorithms [C]// ICEC'96: Proceedings of the 1996 IEEE International Conference on Evolutionary Computation. Washington, DC: IEEE Computer Society, 1996: 61-66.
- [12] HAN K H, KIM J H. Genetic quantum algorithm and its application to combinatorial optimization problem [C]// ICEC 2000: Proceedings of the 2000 IEEE Conference on Evolutionary Computation. Washington, DC: IEEE Computer Society, 2000: 1354-1360.
- [13] 李士勇, 李盼池. 基于实数编码和目标函数梯度的量子遗传算法[J]. 哈尔滨工业大学学报, 2006, 38(8): 1216-1223.
- [14] 王凌. 量子进化算法研究进展[J]. 控制与决策, 2008, 23(12): 1321-1326.
- [15] LI B, ZHUANG Z Q. Genetic algorithm based on the quantum probability representation [C]// Proceedings of the 3rd International Conference on Intelligent Data Engineering and Automated Learning, LNCS 2412. Berlin: Springer-Verlag, 2002: 500-505.