

基于联系数的网格任务动态调度算法

黄德才¹, 龚卫华¹, 张丽君¹, 赵克勤²

(1. 浙江工业大学信息工程学院, 杭州 310032; 2. 诸暨市联系数学研究所, 诸暨 311811)

摘要: 针对传统网格任务调度模型将每个任务的预期执行时间设定为一个确定常量, 无法表达其不确定性的问题, 利用一种新的软计算方法——集对分析联系数来表示并处理网格任务预期执行时间的综合不确定性。提出基于联系数的不确定网格动态调度模型, 在线动态调度算法 OUD_MCT 和批模式动态调度算法 BUD_Surferage。数值仿真结果表明, 该类算法能较好地描述网格任务预期执行时间的动态性和不确定性, 使传统网格调度方法成为其特例, 在动态和不确定网格环境中具有良好的理论意义和实用价值。

关键词: 计算网格; 不确定性; 动态; 任务调度; 调度算法

Grid Task Dynamic Scheduling Algorithm Based on Connective-number

HUANG De-cai¹, GONG Wei-hua¹, ZHANG Li-jun¹, ZHAO Ke-qin²

(1. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310032;

2. Zhuji Institute of Connective Mathematics, Zhuji 311811)

【Abstract】 Aiming at the problem of traditional grid task scheduling model can not express and process the uncertainty existing in grid environment, because it takes the expected time to compute of a task as an constant, and by means of connective number of set pair analysis, this paper presents a new soft computation method to express and process the synthetic uncertainty of task scheduling for computing grid, namely using a connective number to express the expected time to compute of a task. After presenting the uncertain scheduling model based on connective-number, online uncertain dynamic scheduling algorithm OUD_MCT, and batch uncertain dynamic scheduling algorithm BUD_Surferage are presented for the uncertain dynamic computing grid. Numerical value simulation results illustrate that these algorithms can express the dynamics and uncertainty of expected time to compute of tasks in the computing grid environment. These algorithms are the generalization of traditional grid scheduling algorithms, and with high value in theory and application in the uncertain and dynamic grid environment.

【Key words】 computing grid; uncertainty; dynamic; task scheduling; scheduling algorithm

1 概述

任务调度系统是使网格发挥潜在性能和优势的核心系统^[1], 其中, 任务调度模型及其优化算法是网格调度的关键技术。调度算法的好坏将直接影响网格的信息传递效率、资源利用率、网格可靠性和应用程序的执行时间等^[2]。目前已有较多网格任务调度算法^[2-5], 但它们没有考虑网格环境的动态性和不确定性, 导致现有任务调度系统难以在开放、异构和动态的真实网格环境中有效运行。

网格环境中存在大量不确定因素, 例如, 网络流量和资源的当前负载等都是动态变化的, 它们反映出的结果将直接影响任务的执行速度, 导致无法确定任务的预期执行时间。文献[6]提出具有模糊处理时间的人工免疫网格任务调度算法。但由于网格具有广域性、异构性和动态性, 因此网格任务调度中的不确定性不是单纯的模糊或随机不确定性, 而是由随机、模糊、中介、未知和信息不完全等引起的综合不确定性。传统网格任务调度算法难以真实描述和表达上述综合不确定性问题。因此, 文献[7]利用集对分析联系数^[8]研究了网格调度综合不确定性的表达和处理方法, 借鉴传统网格任务调度算法思想, 提出3个基于联系数的不确定网格静态调度算法并分析其优点。针对网格服务需求的动态性, 本文提出基于联系数的网格任务在线动态调度算法 OUD_OLB,

OUD_MET, OUD_MCT(Online Uncertain & Dynamic)和批模式动态调度算法 BUD_Min-min, BUD_Min-max, BUD_Surferage(Batch Uncertain & Dynamic)。

2 联系数全序关系

集对分析是一种新的软计算方法^[8], 联系数是集对分析中的重要概念之一^[9]。联系数的应用方便灵活, 其理论意义在于拓展了数的概念。文献[7]根据计算网格中任务预期执行时间(Expected Time to Compute, ETC)的不确定性特点, 给出联系数定义。

定义 1^[7] 称 $u=a+bi$ 为网格调度问题中表示某个任务预期执行时间的联系数, 其中, a 表示该任务在通常情况下的预期执行时间; b 为网格环境不确定性引起的时间波动; i 在区间 $[-1,1]$ 内不确定取值。特别地, 当 $a=b=0$ 时, 记联系数 $u=0$ 。

定义 2^[7] 设 $u_1=a_1+b_1i, u_2=a_2+b_2i$ 是2个联系数, 则它们

基金项目: 浙江省自然科学基金资助项目(Y105118, Y105109)

作者简介: 黄德才(1958—), 男, 教授、博士、博士生导师, 研究方向: 网格调度, 人工智能, 图像处理, 数据挖掘; 龚卫华, 讲师、博士; 张丽君, 硕士研究生; 赵克勤, 高级工程师

收稿日期: 2008-11-30 **E-mail:** hdc@zjut.edu.cn

的和是一个联系数 $u=a+bi$, 记作 u_1+u_2 , 其中, $a=a_1+a_2$, $b=b_1+b_2$ 。

定义 3^[7] 设 $u_1=a_1+b_1i$, $u_2=a_2+b_2i$ 是 2 个联系数, 则它们的差是一个联系数 $u=a+bi$, 记作 u_1-u_2 , 其中, $a=a_1-a_2$, $b=b_1-b_2$ 。

根据网格调度的实际需要, 文献[7]给出联系数的 2 种全序关系。

定义 4 设 $u_1=a_1+b_1i$, $u_2=a_2+b_2i$ 是 2 个联系数。如果 $a_1 < a_2$ 或 $a_1=a_2$, 且 $b_1 < b_2$, 则称 u_1 小于 u_2 , 记作 $u_1 <_p u_2$ 。

在定义 4 中, “ $<_p$ ” 称为悲观估计下的全序关系。

定理 1^[7] 设 R_c 为所有联系数的集合, 则对于任意 $u_1, u_2 \in R_c$, 以下 3 个结论中有且仅有一个成立: (1) $u_1=u_2$; (2) $u_1 <_p u_2$; (3) $u_1 >_p u_2$ 。

定义 5 设 $u_1=a_1+b_1i$, $u_2=a_2+b_2i$ 是两个联系数。如果 $a_1 < a_2$, 或 $a_1=a_2$, 且 $b_1 > b_2$, 称 u_1 小于 u_2 , 记作 $u_1 <_o u_2$ 。

在定义 5 中, “ $<_o$ ” 称为乐观估计下的全序关系。

定理 2^[7] 设 R_c 为所有联系数的集合, 则对于任意 $u_1, u_2 \in R_c$, 以下 3 个结论中有且仅有一个成立: (1) $u_1=u_2$; (2) $u_1 <_o u_2$; (3) $u_1 >_o u_2$ 。

在本文中, 联系数表示的时间用 “ $<_p$ ” 或 “ $<_o$ ” 进行比较。

3 动态调度算法

3.1 调度问题与参数定义

根据网格任务调度问题的定义^[7], 对本文基于联系数的不确定网格调度问题进行描述。设有 n 项独立任务 $T=\{t_1, t_2, \dots, t_n\}$ 和 m 个计算机资源节点 $M=\{M_1, M_2, \dots, M_m\}$, 任务 t_p 在节点 M_q 上的预期执行时间 $E(p, q)$ 是不确定的, 可以用联系数表示为 $E(p, q)=a_{pq}+b_{pq}i$, 并构成一个 $n \times m$ 的 ETC 矩阵 $(E(p, q))_{n \times m}$ 。任务调度的目标是把 T 中任务安排到 M 中的节点上执行, 使其总执行时间(makespan)尽可能小。对其调度算法涉及的参数定义如下:

(1) M_q , 表示编号为 q 的计算资源(节点), 是该节点上所分配任务的集合, 且按任务分配的先后次序存放。

(2) S_q , 节点 M_q 完成已分配给它的所有任务后, 预期可执行另外一个新任务的开始时间。

(3) $E(p, q)$, 任务 t_p 分配给可用节点 M_q 的预期执行时间, 形成的矩阵称为 ETC 矩阵。

(4) $C(p, q)$, 将新任务 t_p 分配给可用节点 M_q 的预期完成时间, $C(p, q)=S_q+E(p, q)$ 。

(5) CT , 任务与资源匹配形成的二元组 \langle 任务, 资源 \rangle 集合, 又称任务-资源对集合。

(6) $Assigned_q$, 表述在一次循环中, 节点 M_q 是否已经分配一个新任务, 取值为 True 或 False。

3.2 在线模式不确定动态调度算法

算法 1 OUD_OLB

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for M 中的每个节点 M_q

$M_q = \phi$

endfor

for T 中的每个作业 t_p

$S_{min} = \min\{S_1, S_2, \dots, S_m\}$

$M^{min} = \{M_k \mid S_k = S_{min}, k=1, 2, \dots, m\}$

在 M^{min} 中任选一个节点 M_k

$M_q = M_k \cup \{t_p\}, T = T - \{t_p\}$

$S_q = S_q + E(p, q)$

endfor

算法 2 OUD_MET

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for M 中的每个节点 M_q

$M_q = \phi$

endfor

for T 中的每个作业 t_p

$E_{min} = \min\{E(p, 1), E(p, 2), \dots, E(p, m)\}$

$M^{min} = \{M_k \mid E(p, k) = E_{min}, k=1, 2, \dots, m\}$

在 M^{min} 中任选一个节点 M_q

$M_q = M_q \cup \{t_p\}, T = T - \{t_p\}$

endfor

算法 3 OUD_MCT

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for T 中的每个任务 t_p

for 对 M 中的每个节点 M_q

$C(p, q) = 0, M_q = \phi, S_q = 0,$

endfor

endfor

for T 中的每个作业 t_p

for M 中的每个节点 M_q

$C(p, q) = S_q + E(p, q)$

endfor

$C_{min} = \min\{C(p, 1), C(p, 2), \dots, C(p, m)\}$

$M^{min} = \{M_k \mid C(p, k) = C_{min}, k=1, 2, \dots, m\}$

在 M^{min} 中任选一个节点 M_q

$M_q = M_q \cup \{t_p\}, T = T - \{t_p\}$

$S_q = S_q + E(p, q)$

endfor

3.3 批模式不确定动态调度算法

算法 4 BUD_Min-min

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for M 中的每个节点 M_q

$M_q = \phi, S_q = 0$

endfor

Repeat

$CT = \phi$

for T 中每一个作业 t_p

for M 中每一个节点 M_q

$C(p, q) = S_q + E(p, q);$

endfor

找出使任务 t_p 完成时间最短的资源 $M_q, CT = CT \cup \{\langle t_p, M_q \rangle\}$

endfor

$C_{min} = \min\{C(p, q) \mid \langle t_p, M_q \rangle \in CT\}$

$M^{min} = \{M_k \mid C(p, k) = C_{min}$

且 $\langle t_p, M_k \rangle \in CT\}$

在 M^{min} 中任选一个节点 M_k 及其对应的 $\langle t_p, M_k \rangle$

$M_k = M_k \cup \{t_p\}, T = T - \{t_p\}$

$S_k = S_k + E(r, k)$

Until $(T = \phi)$

算法 5 BUD_Min-max

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初

始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for M 中的每个节点 M_q

$M_q = \phi, S_q = 0$

endfor

Repeat

CT = ϕ

for T 中每一个作业 t_p

for M 中每一个节点 M_q

$C(p, q) = S_q + E(p, q);$

endfor

找出使任务 t_p 完成时间最短的资源 $M_q, CT = CT \cup \{ \langle t_p, M_q \rangle \}$

endfor.

$C_{max} = \max \{ C(p, q) \mid \langle t_p, M_q \rangle \in CT \}$

$M^{max} = \{ M_k \mid C(p, k) = C_{max} \text{ 且} \langle t_p, M_k \rangle \in CT \}$

在 M^{max} 中任选节点 M_k 及其对应的 $\langle t_r, M_k \rangle$, 令

$M_k = M_k \cup \{ t_r \}, T = T - \{ t_r \}$

$S_k = S_k + E(r, k)$

Until($T = \phi$)

算法 6 BUD_Sufferage

输入 任务集 T , 节点集 M , 基于联系数的 ETC 矩阵和 S_q 初

始值

输出 任务与资源的映射方案 M_1, M_2, \dots, M_m

for M 中的每个节点 M_q

$S_q = 0, \text{Assigned}_q = \text{False}, M_q = \phi$

endfor

Repeat

for T 中的每个作业 t_p

for M 中每个节点 M_q

$C(p, q) = S_q + E(p, q);$

endfor

for M 中每个节点 M_q

找出使 t_p 完成时间最短的资源 M_r 和对应的完成时间 $C(p, r);$

找出使 t_p 完成时间次短的资源 M_s 和对应的完成时间 $C(p, s);$

$\text{SuffV}_p = C(p, s) - C(p, r)$

endfor

if 任务 t_p 对应的节点 M_q 的 $\text{Assigned}_q = \text{False}$

$M_q = M_q \cup \{ t_p \}, T = T - \{ t_p \};$

$\text{Assigned}_q = \text{True};$

else if t_k 是本次循环已分配给节点 M_q 的任务且 SuffV_p 小于 SuffV_k ,

$M_q = M_q \cup \{ t_p \} - \{ t_k \}$

$T = T \cup \{ t_k \} - \{ t_p \};$

endif.

endfor.

for M 中的每个节点 M_q

$S_q = \sum_{t_r \in M_q} E(p, q)$

$\text{Assigned}_q = \text{False}$

endfor

Until($T = \phi$)

如果算法 4~算法 6 的初始参数 $S_q = 0$, 那么它们就是不确定网络的静态调度算法。

4 数值仿真分析

4.1 数据产生方法

本文仿真实验考察 20 个计算资源组成的网络系统对 40 个~200 个独立任务构成的任务集合的调度情况。

由任务在计算资源上的预期执行时间 ETC, 根据文献[10]方法生成 ETC 矩阵中的 a_{pq} 。根据 a_{pq} 值的大小, 利用均匀分布产生 b_{pq} 。文献[10]方法需要输入 4 个参数, 设控制数据均值变化的参数 $\mu_i = 100, \mu_m = 100$, 另外 2 个参数 V_i 和 V_m 用来控制任务和机器的一致性, 其值越高, 任务和机器的异构性越强, 两者取值为 $V_i = 0.1$ 或 $V_i = 0.6, V_m = 0.1$ 或 $V_m = 0.6$, 使机器和任务分别达到较高或较低的异构性, 可以用来测试异构性的高低对调度是否有大的影响。

4.2 实验结果与性能分析

本文采用悲观估计偏序关系情形下, 任务的平均 makespan 值对改进后的新启发式算法进行综合评价。每个绘图采用 50 次实验的平均值。机器数取 20, 用字母 $m = 20$ 表示。

由于在仿真过程中发现 BUD-Min-max 算法的性能比 BUD-Min-min 算法和 BUD-Sufferage 算法的性能差很多, 因此本文没有对 BUD-Min-max 算法的结果进行比较。

由图 1 和图 2 可以看出, 在非一致 ETC 矩阵条件下, 当任务的异构性较低时, BUD-Min-min 算法的性能优于 BUD-Sufferage 算法(不同于传统方法认为 Sufferage 算法总是优于 Min-min 算法), 但差距不大; 当任务的异构性较高时, BUD-Min-min 算法的性能比 BUD-Sufferage 算法差很多。

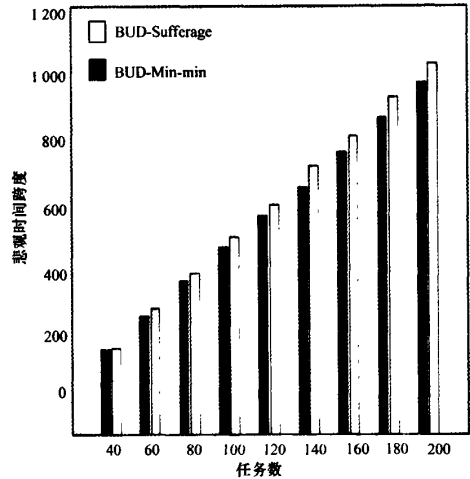


图 1 $m=20, V_i=0.1, V_m=0.1$ 时 2 种批模式算法的性能

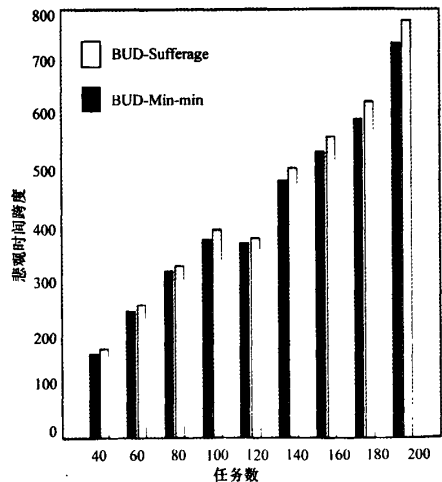


图 2 $m=20, V_i=0.1, V_m=0.6$ 时 2 种批模式算法的性能

3种在线模式调度算法的仿真结果如图3和图4所示。

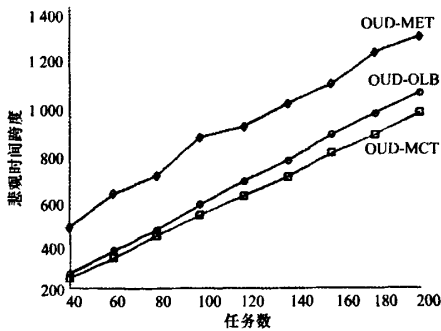


图3 $m=20, V_i=0.1, V_n=0.1$ 时3种在线模式算法的仿真结果

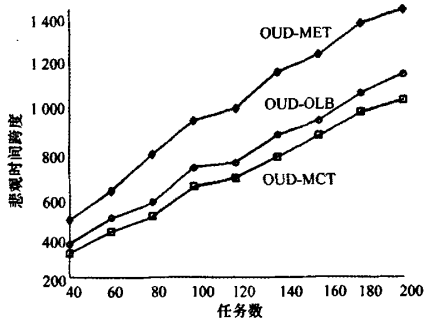


图4 $m=20, V_i=0.6, V_n=0.1$ 时3种在线模式算法的仿真结果

图3和图4表明,在线调度模式下,OUD_MCT算法的

性能最好。而当任务异构性很低时,无论机器异构性如何,OUD-MET算法都优于OUD-OLB算法。

参考文献

- [1] Foster I, Kesselman C. The Anatomy of the Grid: Enabling Scalable Virtual Organizations[J]. International Journal of Supercomputer Applications, 2001, 15(3): 200-222.
- [2] 罗红, 慕德俊, 邓智群, 等. 网格计算中任务调度研究综述[J]. 计算机应用研究, 2005, 22(5): 16-19.
- [3] 陈志刚, 刘安丰, 熊策, 等. 一种有效负载均衡的网格Web服务体系结构模型[J]. 计算机学报, 2005, 28(4): 458-466.
- [4] 陈宏伟, 王汝传. 费用-时间优化的网格有向无环图调度算法[J]. 电子学报, 2005, 33(8): 1375-1380.
- [5] 张伟哲, 刘欣然, 云晓春, 等. 信任驱动网格作业调度算法[J]. 通信学报, 2006, 27(2): 73-79.
- [6] 李季, 钟将, 吴中福. 具有模糊处理时间的网格任务调度免疫算法[J]. 计算机科学, 2006, 33(2): 35-38.
- [7] 黄德才, 张丽君, 赵克勤. 基于a+bi型联系系数的不确定网格静态调度算法[J]. 计算机科学, 2007, 34(8): 126-129.
- [8] 蒋云良, 徐从富. 集对分析理论及其应用研究进展[J]. 计算机科学, 2006, 33(1): 205-209.
- [9] 赵克勤. 集对分析及其初步应用[M]. 1版. 杭州: 浙江科学技术出版社, 2000.
- [10] Ali S, Siegel H J, Maheswaran M, et al. Representing Task and Machine Heterogeneities for Heterogeneous Computing Systems[J]. Tamkang Journal of Science and Engr., 2000, 3(3): 195-207.

编辑 陈晖

(上接第111页)

4 系统应用

本文实现的网格信息归档服务可以方便地部署到基于Globus的网格门户中,并支持扩展应用。目前已部署在福州大学教育部空间数据挖掘与信息共享重点实验室开发的GeoKS-Grid知识网格门户中。该归档服务配合Ganglia, MDS等监控系统,实现了历史数据的存储。基于该数据集,采用HyperbolicTree可视化技术,形象地表达了网格内所有已启动节点的连接关系,并能查询各节点的历史数据和瞬时状态数据。网格信息归档服务应用于GeoKS-Grid门户时的界面见图3。

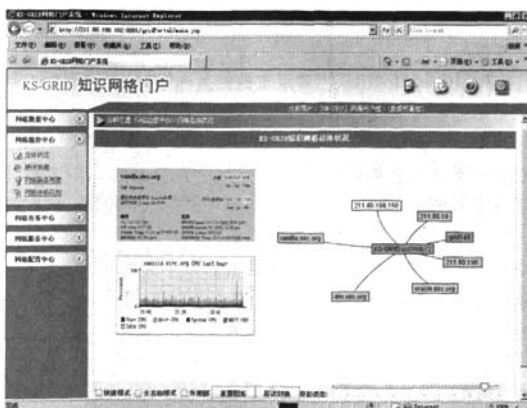


图3 网格信息归档服务应用于GeoKS-Grid门户时的界面

5 结束语

网格运行管理和高级应用服务的开发需要丰富的网格历史状态信息。网格状态信息归档服务的实现满足了网格状态历史信息的查询需求,并能较好地支持基于其上的应用系统的开发。在本文工作的基础上,可以开发一些高级应用服务,例如数据查询接口、历史数据和时序数据的图形表达工具等,以丰富该服务的功能。

参考文献

- [1] Foster I. Globus Toolkit Version 4: Software for Service-oriented Systems[C]//Proc. of IFIP International Conference on Network and Parallel Computing. [S. l.]: Springer-Verlag, 2006.
- [2] Zhang Xuehai, Schopf J. Performance Analysis of the Globus Toolkit Monitoring and Discovery Service[C]//Proceedings of the International Workshop on Middleware Performance. [S. l.]: IEEE Press, 2004.
- [3] Czajkowski K, Kesselman C, Fitzgerald S, et al. Grid Information Services for Distributed Resource Sharing[C]//Proceedings of the 10th IEEE International Symposium. San Francisco, California, USA: [s. n.], 2001.
- [4] Humphrey M, Wasson G, Jackson K, et al. State and Events for Web Services: A Comparison of Five WS-resource Framework and WS-notification Implementations[C]//Proc. of the 4th IEEE International Symposium on High Performance Distributed Computing. [S. l.]: IEEE Press, 2005.
- [5] Andreozzi S. Information Models for the Grid Resources[Z]. 2003.

编辑 陈晖

基于联系数的网格任务动态调度算法

作者: [黄德才](#), [龚卫华](#), [张丽君](#), [赵克勤](#), [HUANG De-cai](#), [GONG Wei-hua](#), [ZHANG Li-jun](#), [ZHAO Ke-qin](#)

作者单位: [黄德才, 龚卫华, 张丽君, HUANG De-cai, GONG Wei-hua, ZHANG Li-jun \(浙江工业大学信息工程学院, 杭州, 310032\)](#), [赵克勤, ZHAO Ke-qin \(诸暨市联系数学研究所, 诸暨, 311811\)](#)

刊名: [计算机工程](#) **ISTIC|PKU**

英文刊名: [COMPUTER ENGINEERING](#)

年, 卷(期): 2009, 35(8)

被引用次数: 0次

参考文献(10条)

1. Foster I, Kesselman C [The Anatomy of the Grid: Enabling Scalable Virtual Organizations](#) 2001(03)
2. 罗红, 慕德俊, 邓智群 [网格计算中任务调度研究综述](#) [期刊论文]-[计算机应用研究](#) 2005(05)
3. 陈志刚, 刘安丰, 熊策 [一种有效负载均衡的网格Web服务体系结构模型](#) [期刊论文]-[计算机学报](#) 2005(04)
4. 陈宏伟, 王汝传 [费用-时间优化的网格有向无环图调度算法](#) [期刊论文]-[电子学报](#) 2005(08)
5. 张伟哲, 刘欣然, 云晓春 [信任驱动网格作业调度算法](#) [期刊论文]-[通信学报](#) 2006(02)
6. 李季, 钟将, 吴中福 [具有模糊处理时间的网格任务调度免疫算法](#) [期刊论文]-[计算机科学](#) 2006(02)
7. 黄德才, 张丽君, 赵克勤 [基于a+bi型联系数的不确定网格静态调度算法](#) [期刊论文]-[计算机科学](#) 2007(08)
8. 蒋云良, 徐从富 [集对分析理论及其应用研究进展](#) [期刊论文]-[计算机科学](#) 2006(01)
9. 赵克勤 [集对分析及其初步应用](#) 2000
10. Ali S, Siegel H J, Maheswaran M [Representing Task and Machine Heterogeneities for Heterogeneous Computing Systems](#) 2000(03)

相似文献(10条)

1. 学位论文 黄昌勤 计算网格中任务管理的若干问题研究 2005

计算网格是一种正在兴起的计算基础设施, 它能协调地理上分布的各类资源为用户提供有效的一致访问, 以满足虚拟组织VOs (Virtual Organizations) 的需要. 计算网络的巨大潜能正吸引着众多的研究者和工程人员研究其技术以支持跨域的协同工作. 然而, 网络系统呈现的性能因为资源竞争和任务不确定性等因素而高度动态变化, 因此, 网络资源调度和任务管理是计算网格中一个关键性的研究课题. 在网格环境中, 任务从提交给网络系统至任务结果处理过程完成, 都一直处于网络任务管理系统的操控之下. 由于网格的大规模、异构、动态、分布和自治等特性, 使得任务管理变得更加复杂和难以有效实现. 本研究以大规模科学和工程计算为背景, 以计算网格为基础环境, 对网络任务管理中的诸多问题进行了研究. 文中所述研究皆得到了国家杰出青年基金项目(60225009)和国家863/CIMS主题资助项目(2002AA414070)的资助. 本文所提出的任务管理系统是MASSIVE (Multidisciplinary Application-oriented SIMulation and Visualization Environment) 工程的一部分, 后者旨在构建面向多学科应用的模拟与可视化网络环境. 本研究主要是从任务调度、网络不确定性处理及QoS操纵、任务授权三方面展开的. 在任务调度的研究中, 一个双组件双队列分布式调度模型DC³>SM (a Dual-Component and Dual-Queue Distributed Schedule Model) 被提出, 基于此模型, 研究了两种调度策略用以来提高网格计算性能. 为了更好的应对网络环境的不确定性, 对DC³>SM调度模型进行了扩展以增加专门不确定性处理辅助部件. 以扩展模型中各个模块的协作为基础, 通过对全局调度器调度策略的修订来解决网络的不确定性; 任务授权是网络基础设施中的重要挑战. 在考虑社区策略的前提下, 提出了并行化子任务级授权服务PSAS (a Parallelized Subtask-level Authorization Service) 体系结构以满足权限的最小化原则, 同时给出了柔性任务管理和上下文感知的动态授权方法.

2. 期刊论文 黄德才, 张丽君, 赵克勤, HUANG De-Cai, ZHANG Li-Jun, ZHAO Ke-Qin 基于a+bi型联系数的不确定网格静态调度算法 - [计算机科学](#) 2007, 34(8)

任务调度算法是计算网格任务管理系统中的核心问题. 由于网格环境中存在大量的不确定因素, 导致传统网格调度算法和调度系统不能在开放、异构和动态的真实网格环境中有效运行. 利用一种新的软计算方法—集对分析联系数研究和处理网格调度中的综合不确定性问题. 在简单介绍集对分析概念和应用情况基础上, 引入联系数概念、运算规律和全序关系, 分别提出了基于联系数的不确定网格静态调度算法CBU-Min-min、CBU-Max-min和CBU-Surferage, 并进行了数值仿真实验研究. 理论和实验研究表明, 这些算法能较好地描述网格任务预期执行时间的动态性和不确定性, 并使传统网格调度方法成为其特例, 在动态和不确定网格环境中有良好的理论和实际应用价值.

3. 学位论文 单世民 基于网格和密度的数据流聚类方法研究 2006

随着硬件技术的不断发展, 人们遇到了大量无法利用数据库进行存储的海量数据. 这些数据数量非常巨大, 并且产生速度很快. 为了对这些数据进行有效处理, 人们提出了数据流数据模型. 作为数据挖掘在新环境的延伸, 面向数据流的数据挖掘 (简称数据流挖掘) 问题已成为当前国内外研究的焦点, 而数据流聚类方法是数据流挖掘的一个重要研究方向.

本文的研究目标是以研究传统聚类方法为基础, 通过对传统方法的改进, 实现对数据流的聚类处理. 通过研究, 发现基于网格和密度的聚类方法具有很多适用于处理数据流的特征, 有利于实现对数据流的聚类处理. 因此, 本文在对基于网格和密度的传统聚类方法进行研究与改进的基础上, 从聚类过程所处理数据集合的动态性角度出发, 将基于网格和密度的数据流聚类方法分类为静态方法和动态方法, 对其进行了一系列研究.

围绕着数据流聚类问题, 论文主要做了以下四方面的理论研究及应用工作:

1. 对基于网格和密度的数据流静态聚类方法进行了讨论, 通过对传统的基于网格和密度的聚类方法进行分析与改进, 提出了一种新的网格单元密度计算方法. 研究发现, 已有方法基本使用数据点计数方式计算网格单元的密度. 这种方法会造成数据点对其周围空间影响信息 (influence) 的部分丢失, 容易导致同属一类的相邻数据点被分配到不同的数据类中. 针对此问题, 提出了“贡献度”概念. “贡献度”即是指在网格化的特征空间中, 数据点对相邻网格单元 (即数据点的周围空间) 的影响程度. 在此基础上, 提出了一种新的网格单元密度计算方法. 实验结果证明, 与利用数据对象个数计算网

格单元密度的方法相比,这种方法能够有效减少数据点对周围空间影响信息的丢失。

2. 针对已有的基于网格和密度的聚类方法在稠密单元判定方式以及聚类生成过程方面的不足进行了改进,最终提出了一种新的基于网格和密度的微粒混合聚类方法。现有方法基本使用单一阈值来判定稠密网格单元。这种方法限制了对聚类数据与噪音数据进行区分的能力。同时,已有方法都将首个遇到的稠密单元作为聚类生成过程的起始点。这种操作对生成的数据类缺乏选择性,使数据类的生成顺序具有不确定性,进而影响对结果的可控性。针对这些问题,使用了新的参数一“核心单元下限”。此参数限定了聚类中密度极大值的最小允许取值,对能够生成的聚类进行了限制,增强了区分聚类数据和噪音数据的能力。同时,通过将微粒群算法引入聚类过程,使得数据类能够根据本身的密度极大值有序生成,解决了聚类生成顺序的不确定性问题。最终,提出了基于网格和密度的微粒混合聚类方法(CGDP)。

3. 讨论了动态的网格空间环境下对聚类进行追踪的方法,提出了动态环境下改进的自适应微粒群算法。通过将网格单元密度转化评估函数取值的方式,把对聚类密度极值的搜索问题转化为优化问题,然后利用微粒群算法解决此优化问题。针对数据流的动态性特征,分析了已有的动态环境下微粒群算法,发现已有方法的环境变化检测能力仍然存在不足,具体表现为微粒种群容易停滞在动态环境中静止的局部极点,进而失去对其他区域中全局极点的追踪能力。针对此问题,使用“活性因数”概念及分布式处理模式,解决了微粒种群在动态环境中的停滞问题。进一步的,提出了动态环境下改进的自适应微粒群算法(IAPSO)。实验证明,该方法能够适应更多类型的复杂动态环境,具有更为普遍的实际意义,为论文对数据流聚类方法的研究提供了技术支持。

4. 以入侵检测为代表的具数据流环境为研究对象,对基于网格和密度的数据流动态聚类方法进行了研究,提出了数据流中孤立点识别方法。在此类数据流环境中,可以认为只存在一个主要聚类,并且与其他数据相比,属于主要聚类的数据具有相对更高的密集程度。这种环境的基本需求就是判断新出现的数据对象是否属于主要聚类。针对这种情况,首先扩展了“孤立点”的含义,将不属于主要聚类的数据对象定义为相对于主要聚类而言的“孤立点”。然后,将本文的前三项工作成果进行综合,提出了一种面向数据流的动态聚类方法—数据流中孤立点识别方法(ODODS)。设计并实现了数据流分析原型系统,并在此基础上,以入侵检测的标准测试数据为基准进行实验,以满足具体数据流环境的应用需求为目标,通过与以往基于聚类的入侵检测方法进行结果对比,说明了ODODS方法的有效性。

4. 学位论文 [任姝娜 基于Gridsim的网络资源调度算法及其应用的研究](#) 2004

网络计算是近年来得到快速发展的广域网络计算技术。研究人员试图将很大范围上地理分布的异构计算机系统集合在一起形成一个大规模的计算平台,我们称之为网格(Grid)。通过这种平台,用户能够方便、快捷地使用自己需要的资源而不必了解其中的细节。网络系统提供给用户的除了廉价、高效的高性能计算以外,还包括合作存取各种数据信息,广域多媒体应用等等。

然而,由于网格是个新的研究领域,网格环境相对于一般网络环境有着更为复杂的特征,如存在多管理域和站点自治,系统的动态性、异构性和通信延迟的不确定性更高,硬件和软件两个层次上都存在异构性等等。因此,实现有效的网络计算还有很多需要解决的问题,具体包括资源调度和管理、系统安全、编程模式、性能评测和数据存取等。其中,在网格环境里如何有效地管理资源和调度计算是影响网络计算是否成功的最重要因素之一。由于资源在广域网上的分布性、异构性以及存在着不同的存取和花费模式,使得网格环境下的资源管理变得十分复杂和具有挑战性。

本文试图从网格资源的特点出发,寻求一种更加合理的网络资源管理模式和更加有效的调度策略。作者参阅了国内外的大量文献,在研究目前的网络资源管理模型的基础上,提出了基于经济的网格市场模型并且实现了受任务完成时间和费用约束的贪婪算法,取得了一定的研究成果。

本文中作者的贡献是:(1)针对目前国际上对网络资源管理的研究所采取的普通的方法,深入分析了现有网络资源管理模型及计算网络系统的特点,提出基于经济的网格市场模型的资源调度管理方法。

(2)借用市场、拍卖等经济学模型描述了基于经济的网格市场模型的资源调度管理的内部机制和行为。

(3)研究网格计算环境下一组相互独立的计算任务(Meta-Task)的资源映射策略,分析了常用的资源调度算法,并且比较了它们的优缺点和适用情况。

(4)在基于经济模型的网格模拟器Gridsim环境下,实现了受任务完成时间以及费用约束的贪婪算法,在追求较小的任务完成时间的同时兼顾任务的服务质量(QoS)需求,使之能够更好地支持大规模用户使用网格资源。

5. 会议论文 [黄德才.刘端阳.张丽君.赵克勤 不确定网络调度问题的联系数学建模方法初探法](#) 2007

任务调度算法是计算网络任务管理系统中的核心技术问题。由于网格环境中存在大量的不确定因素,导致传统网络调度算法和调度系统不能在开放、异构和动态的真实网格环境中有效运行。利用一种新的软计算方法—集对分析联系数研究和处理网络调度中的综合不确定性问题。在简单介绍分析概念基础上,引入联系数全序关系,提出了一个基于联系数的不确定网络静态调度算法CBU_Min-min,并进行了数值实例计算。理论和实例计算表明,这种模型能较好地描述网络任务预期执行时间的动态性和不确定性,并使传统网络调度方法成为其特例,在动态和不确定网络环境中有良好的理论和实际应用价值。

6. 学位论文 [吴娟 基于策略域的分布式访问控制模型](#) 2009

随着越来越多的分布式系统的出现,如P2P系统、跨多个域的计算网格、Web服务组合,软件系统的形态发生了根本性地变化。作为软件系统安全保障的访问控制系统也需要重新的设计和实现。本文针对现有访问控制模型在分布式环境中表现出的不足,以对各种分布式访问控制系统的定性研究为基础,设计并实现了一个分布式访问控制模型,取得了以下重要成果:

1. 对典型的分布式访问控制系统进行了定性研究。首先从权限委派方式、身份鉴别机制、状态信息的维护、决策模型、信任度评估、实现与方案的一致性这6个方面进行了性质分析。然后以这几个指标为依据,对具体的访问控制系统做了量化。基于这些研究对这一方向的研究现状做出了总结,并定下了本文研究工作的目标和基调。

2. 对分布式系统做了抽象,提出了策略域的概念,并且基于策略域,提出了一个分布式访问控制模型PDAC。对于访问控制策略,文章首次将能力和契约结合在一起使用,能力从客观上反映了节点提供的操作,而对这些操作的约束则由策略域实时地给出。这样,PDAC模型就具备了动态授权的特征。对于决策模型,策略域既可以根据本地信息做出独立决策,也可以联合其他策略域根据全局信息做出合意决策。因此,PDAC模型具备分布决策的特征。另外,PDAC模型支持权限的委派,并且当委派权限的策略域对权限进行修改后,策略域能够记录这种改变,使得我们能够以更加细粒度的方式来管理授权。

3. 以联想广播为基础,实现了策略域的协同工作模型。使用该模型,策略域可以通过指定状态来选择希望与之协作的其他策略域,也可以通过设置自己的本地状态来获得与其他策略域协作的机会。基于联想广播的协同工作模型能够在—组策略域中动态地定义多个协同子集,满足分布式环境的动态协作要求。以这个协同模型为基础,为PDAC模型制定了服务发现机制和领域查询机制。

4. 基于主观逻辑方法,在信任评估中引入不确定性因素,确立了动态信任评估机制。实验结果表明,基于主观逻辑的信任评估能够更好地反映分布式环境的不确定性,对于促进节点间的协作有着积极的作用。考虑到现实世界的信任具有随时间衰减的特征,在信任评估模型中引入了老化机制。信任评估的老化能够客观地反映节点间的信任关系,对节点的行为起到了很好地约束作用。

5. 结合BAN逻辑和时序逻辑,提出了策略描述语言PDPL,给出了形式化的语法规则和语义解释。进而,以PDAC模型作为形式系统的一个语义模型,定义了形式系统的公理和推理规则。以这些工作为基础,对PDAC模型的握手、委派、协作、决策等性质进行了可靠性论证。

总的来说,本文提出的PDAC模型是对访问控制模型的一种新的解决思路,本文的研究成果推动了分布式访问控制模型研究的发展,对构建更加安全可靠的分布式系统具有重要的实践指导意义。

7. 期刊论文 [殷锋.何先波.刘韬.YIN Feng.HE Xian-bo.LIU Tao 基于自适应模糊-PID反馈模型的网格调度技术](#)—[计算机学报](#)2009,36(7)

服务调度技术是计算网络任务管理系统中的核心问题。为适应网络服务部署的需要,提出了一种基于自适应“模糊-PID”反馈控制模型的Agent技术,该技术融合了模糊理论与PID(Proportionment-Integral-Differential coefficient)技术的优点。实验证明本技术能解决网络服务部署中的动态性和不确定性,并充分发挥网络服务的虚拟执行功能,可在网络带宽效率、延时和可靠性等方面做出更好的权衡。

8. 学位论文 [王铁强 实际地形下风场CFD模拟与参数分析](#) 2009

利用计算流体力学(Computational Fluid Dynamics, CFD)模型模拟近地面风场是风能利用的技术发展趋势。由于风场CFD模拟的复杂性,现阶段在面向风能利用的风场模拟中应用还比较少。本文的主要目的是通过风场模拟实验,研究风场CFD模拟的可行性和模拟精度。

本文选择了具有较完整现场测风数据的苏格兰Askervein山地作为模拟计算场址,解决了真实地形的建模问题,并对各种近地面风场计算网格的生成

方法进行了研究。在此基础之上, 本文通过实验从首层网格高度和网格高度增长率方面考察了网格精度对模拟结果的影响; 比较了多种湍流模型在实际地形下的近地面风场模拟中的表现, 并总结出具有较高模拟精度、适用于近地面风场模拟计算的湍流模型; 研究了不同离散格式对模拟结果的影响, 并总结了各种离散格式的适用范围; 分析了不同的边界条件如地表粗糙度条件对模拟结果的影响, 并提出了适用于近地面风场模拟计算的边界条件设置方案。

本文建立了利用CFD模型进行实际地形下近地面风场模拟计算的总体方案及流程, 提出了模拟计算的自动化方案, 并得到了较好的应用效果, 一方面证明了利用CFD模型进行风场模拟的可行性, 另一方面也表明利用CFD模型进行风场模拟具有较高的准确性, 能降低风场模拟的不确定性。

9. 期刊论文 [张辉亚, 张煜盛, 肖合林, 徐波 提高喷雾模拟精度的算法 - 科学通报](#)2007, 52(5)

燃油喷雾是直喷式发动机燃烧的关键过程, 喷雾模拟的精度决定了燃烧计算的可靠性。然而, 在KIVA和商业CFD代码中, 传统喷雾模拟技术对网格精度非常敏感。因而, 预测的发动机性能和排放依赖于计算网格。导致这个问题的两个主要原因是液滴碰撞算法和气泡相间耦合。为了提高喷雾模拟精度, 采用交错网格液滴碰撞(CMC)算法和气相速度插值算法对原始KIVA代码进行了修正。在定容燃烧室和直喷式柴油机条件下, 分别从喷雾结构、预测平均液滴尺寸和喷雾贯穿距三个方面检验了改进KIVA代码对喷雾模拟精度的改善。结果表明, 网格依赖性显著降低。通过这些改进, 喷雾结构的失真现象消失。在定容燃烧室计算中, 预测的平均液滴尺寸的不确定性从30 μm 减小到5 μm ; 在发动机模拟中, 这种不确定性进一步减小到2 μm 。在中等和精细网格条件下, 发动机模拟预测的喷雾贯穿距也获得了较好的一致性。

10. 学位论文 [马晨明 基于多QoS的网格资源管理与任务调度算法研究](#) 2008

网络是高性能计算和信息服务的战略性基础设施, 而网络技术已成为下一代互联网应用的关键技术。高速网络的发展将分散的、异构的计算资源有机地整合到一起形成计算网格, 为解决科学和工程领域一些大规模计算问题提供了理想的平台。由于网络所具有的广域性、动态性、异构性等特点, 如何对资源进行管理及任务调度以满足用户的需求是一个极具挑战性的问题。

随着网络的飞速发展, 服务质量(QoS)成为网格资源管理与任务调度过程中必须要考虑的一个重要因素。在这种情况下, 本文通过对网格QoS框架的研究, 将多QoS的模糊属性融合到资源的管理与任务调度算法中, 使网格任务在不确定环境下按时执行, 本文的主要工作包括以下几部分:

1. 首先介绍了网格计算的概念和目的、网格QoS的体系结构并同时指出了基于QoS任务调度的重要性; 然后在此基础上对网格技术与调度算法的研究现状进行了详细的阐述和总结。

2. 针对用户QoS需求的模糊性和不确定性, 提出基于D—S理论的区间值模糊多QoS测量方法, 将用户提交的区间QoS需求有效地进行融合, 形成优先队列, 仿真结果表明, 该方法可以以区间数的形式满足用户的多QoS的任务调度需求。

3. 为了能在任务调度中更快地发现资源, 对网格资源的合理管理显得尤为重要。通过对资源的QoS属性进行划分, 利用数据挖掘中改进的DBSCAN聚类算法将网格环境中的各异构资源进行资源绑定是一种有效的资源管理方法。

4. 基于网格环境中的多QoS因素, 改进了现有的网格调度算法, 提出了以makespan值作为优先目标的网格调度新算法, 仿真验证实算法在不同异构环境下都有很好的效果。

5. 不确定网格环境中通常将服务质量的高低作为任务的优先调度目标, 借鉴人类社会学中的人与人之间的信任模型, 考虑网格节点间的服务质量信任需求, 提出了基于QoS的网格关联任务调度算法。算法考虑任务之间的通信与执行的先后性, 更符合实际的网格环境, 该算法以较小的时间跨度为代价, 取得了较好的服务质量性能。

最后, 对本文的研究工作进行了总结, 对目前研究内容存在的问题进行了分析, 并提出了进一步的展望。

本文链接: http://d.g.wanfangdata.com.cn/Periodical_jsjgc200908038.aspx

授权使用: 浙江工业大学图书馆(wfzjgydx), 授权号: e8b3e486-7634-43b1-9c03-9e0000ab03cb

下载时间: 2010年9月29日