

基于负载均衡的任务调度改进算法

韩 云¹, 于 炯¹, 张 伟², 王命全¹

(1 新疆大学 信息科学与工程学院, 新疆 乌鲁木齐 830046; 2 河南省科学院 应用物理研究所, 河南 郑州 450008)

摘 要: 网格环境下, 任务和资源综合性能的预测与评估是动态环境下资源与任务合理匹配的前提, 也是保证网格服务质量的基础. 文中基于用户稳定性、资源性能、网络带宽以及距离因子对可用资源的整体性能和状态进行预测和评估, 并利用资源和任务的性能预测值设计了一种新的调度策略. 仿真实验验证了该算法的有效性.

关键词: 网格计算; 资源性能; 服务质量; 负载均衡

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2010)08-0201-04

An Task Scheduling Algorithm Based on Load Balance

HAN Yun¹, YU Jiong¹, ZHANG Wei², WANG Ming-qun¹

(1 School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China;

2 Henan Academy of Sciences, Institute of Applied Physics, Henan, Zhengzhou 450008, China)

Abstract: In the grid environment, performance prediction and assessment is the premise of a reasonable match in a dynamic environment, and ensure the quality of service. It is to predict and evaluate the overall performance of resources available based on the user's stability, resource properties, network bandwidth and the distance. And using performance prediction of the resources and task design a kind of new scheduling strategy. The simulation results verify the effectiveness of the proposed algorithm.

Key words: grid; performance; QoS; balance

1 引言

在网格动态环境下, 任务调度的研究一直是学术界研究的重点, 网格系统最大化吞吐量和最小化任务执行时间是网格计算^[1]研究的目标. 由于网格环境的开放性、动态性、异构性等特点, 网格服务不可能提供一种固定的、统一的映射策略, 事实证明这种映射是 NP 完全难题.

目前, 针对独立任务的映射策略可分为静态调度算法和动态调度算法. 由于网格资源的动态性, 采用静态调度方法不能真实地反映网格环境下的实际调度情况. 因此, 静态调度方法在动态的网格环境下几乎没有实际意义. 相比之下, 动态调度算法的参数更新较快, 在复杂环境下操作性能良好, 算法灵活, 更适合于变化的网格环境.

网格系统含有大量为不同组织服务且异构的资源, 各个组织对其内部资源和任务的管理机制不同, 资源和任务的调度策略、费用和调度目标都不尽相同. 因此, 为了满足用户任务的不同需求, 最好的办法是建立基于多维服务质量的管理和调度模型, 设计基于服务质量的调度算法及策略考察任务和资源的 QoS, 完成任务和资源的合理映射.

QoS 指导的 min-min 算法^[2]简单地将任务分为两类, 一类是有 QoS 要求的, 另一类是没有 QoS 要求的, 然后分别对这两类任务集合用 min-min 算法进行调度. 在这个算法中, 虽然考虑了任务的服务质量要求, 但是该算法只是简单地将 QoS 作为一维因素考虑进去, 不符合现实网格环境下多维 QoS 要求的实际需求. 在复杂的网格环境下, 设计全面考察任务和资源综合需求和性能的算法势在必行. 目

前出现的算法虽然也考虑了三维到五维的 QoS 需求,但算法总体效果不是很好.一些传统算法^[3]只是对现有算法的简单改进,在调度效果上并没有质的飞跃.

因此,文中借鉴安全驱动下的信任调度思想^[4],根据任务和资源的 QoS 综合需求,把任务的 QoS 分为两类,并将资源负载时间作为衡量资源负载均衡的重要因素,设计了一种新的调度算法 TR-BS(task resource-balance schedule)算法.该算法根据任务的需求,从资源列中找到一个较优的资源来执行它.

文中要解决的问题是如何将 n 个相互独立的任务分配到网络系统中 m 个计算资源上.网络系统的广泛应用,要求用户在提交任务时给出 QoS 需求参数.网络环境下任务的服务质量有很多种,比如,计算密集型任务,这类任务对计算要求比较精密,即对资源本身的计算性能要求较高;运输型任务,这类任务对网络带宽要求较高,否则会在传输中造成很大的时间延迟.常见的任务 QoS 包括:cpu 处理能力、网络带宽、可靠性、安全性、准确性等.

网络任务服务质量的评估方法可以简单地分为两种^[5]. (1)直接通过量化或者未量化的任务属性值大小进行比较.这种策略一般适合一维 QoS 的评估; (2)通过对任务的各个属性值进行加权计算出任务的综合性能,然后进行比较.其主要思想是:将任务的各维 QoS 参数分别量化,然后求该任务的综合性能.这种方法适合网络任务具有多维 QoS 属性的情况.文中就是根据第 2 种方法对任务进行综合性能的评估.

2 任务 QoS 分类及需求描述

根据不同的标准,网络任务的划分方法也不同.文中根据网络任务需求参数的大小和任务需求质量的高低将任务的 QoS 分为两类: (1)积极 QoS.指参数值越大质量越高的任务需求,如可靠性、带宽、存储能力等; (2)消极 QoS.指参数值越大质量越小的任务需求,如响应时间、服务费用和任务执行时间等.在计算任务 QoS 需求时,将这些 QoS 参数简单的加权平均意义不大,这样两类 QoS 参数大小表示任务需求完全相反的参数在一个评估公式中也是不能实现的.因此必须把这两类 QoS 分别处理,最后再统一起来才是明智的选择.

用 $T = \{T_1, T_2, \dots, T_n\}$ 表示 n 个任务的集合, T_i 表示第 i 个任务.假设每一个任务有 k 维 QoS

需求,为了能形象描述 QoS 参数,将每一个任务的 k 维 QoS 参数用一个一维向量来表示:

$$q^{task} = (q_1^{task}, q_2^{task}, \dots, q_k^{task}) \quad (1)$$

从式(1)可以看出,对原始度量值进行加权平均没有意义,因为式中的参数既有积极参数也有消极参数.为了更合理地评价网络任务服务质量的高低,对式(1)中各维 QoS 参数值进行统一量化.任务需求较高的被量化为 1,需求较低的被量化为 0,因此,任务的需求被量化到区间 $[0, 1]$ 上.量化的公式具体表达式如下:

$$\text{Val}(q_{ij}) = \frac{q_{ij} - \min}{\max - \min} \quad (2)$$

式(2)对网络任务的积极 QoS 参数进行量化,将参数大小不一的数值量化到区间 $[0, 1]$ 上,量化后的参数大小和任务服务质量需求的高低成正比例.

$$\text{Val}(q_{ij}) = \frac{\max - q_{ij}}{\max - \min} \quad (3)$$

式(3)对网络任务的消极 QoS 参数进行量化,量化后的数值也统一到区间 $[0, 1]$ 上,量化后的参数大小和任务服务质量需求的高低也成正比例.式(2)与(3)是建立在 $\text{Max} \neq \text{Min}$ 基础之上的,当 $\text{Max} = \text{Min}$ 时, $\text{Val}(q_{ij}) = 1$. 其中, $\text{Max} = \text{Max}(q_{ij})$; $\text{Min} = \text{Min}(q_{ij})$. $1 \leq i \leq n, 1 \leq j \leq k, n$ 为候选网络任务的个数, k 为 QoS 参数的总维数.

网络任务 QoS 综合需求表示如下:

$$\text{GT} = \sum_{j=1}^k W_j * \text{Val}(q_{ij}) \quad (4)$$

其中 W_j 为第 j 维 QoS 的权值, $0 \leq W_j \leq 1$ 且 $\sum_{j=1}^k W_j = 1$.

由式(4)可以计算出每一个网络任务的综合参数值 GT. GT 的值越趋近 1,说明该网络任务的相对服务质量需求越高.

3 资源性能的参数衡量

网络系统中的资源参数是时时刻刻都在变化的,这样的环境给调度带来了很大的困难.为了简化调度算法应付资源动态性复杂的处理过程,这里提出了资源快照的概念,即在批调度时刻,节点服务端通过查询网络资源^[6]服务器,获得网络系统当前可用资源的列表.从网络资源的动态考虑,这种实时系统资源快照有助于调度算法适应网络系统的资源动态变化特性.在一次调度中,调度策略获得的当前系统资源快照是一个相对静止的可用资源集合.

假设网格系统可用的有效资源集合为 R , 且资源节点之间是全连通的. 用 $R = \{R_1, R_2, \dots, R_m\}$ 表示 m 个计算资源的集合, R_j 表示第 j 个计算资源.

下面介绍几个影响网格资源性能的主要因素:

定义 1 资源稳定性参数 S : 该参数值能相对准确地反映资源的在线情况. S 值由用户在线时间长短决定, 它是指网格资源提供者从上线到下线的在线时间平均值. 该参数值可以在网格系统的服务中心获取, 经过归一化处理后, 取值范围是 $[0, 1]$.

定义 2 资源处理器性能参数 P : 该参数反映了资源可以提供的处理能力, 一般主要指 cpu 的处理速度, 由网格资源提供者在网格系统中注册时给出. 归一化之后的取值范围是 $[0, 1]$.

定义 3 资源带宽参数 b : 该参数是指资源能够提供的传输带宽, 带宽的大小将直接影响任务的传输时间, 特别是对于传输型的任务来说, 直接影响网络的计算时间, 因为网格任务的完成时间等于计算时间与传输时间之和. 资源的传输带宽参数值由网格资源提供者在系统中注册时给出, 并经过归一化, 取值范围是 $[0, 1]$.

定义 4 负载参数 q : 资源的负载情况直接影响到任务的完成时间跨度和整个网格系统的工作效率. 资源负载情况由当前资源负载的任务执行时间决定, 时间越长, 负载越重, 时间越短, 负载越轻. 为了方便计算, 参数 q 的值处理后取值范围是 $[3, 4]$.

定义 5 距离参数 dis :

$$dis = d * w_d \quad (5)$$

在式(5)中, d 表示任务与资源之间的网络距离, 可以利用网络坐标计算得到. d 的取值范围是 $[0, 1]$, w_d 为性能调整权重因子.

定义 6 引力因子 g :

$$g = g(s, p, b) \\ = g(s * W_s + p * W_p + b * W_b) \quad (6)$$

其中, W_s 、 W_p 、 W_b 为权重因子, 分别用于用户稳定性、cpu 处理速度和网络带宽的性能权重调整. g 值越大, 说明资源越稳定, 资源的处理器性能越高, 拥有的网络带宽越大. 每一个节点的引力因子 g 将随 IP 和端口号一起记录在网格代理节点的服务器中.

综合以上因素, 资源综合性能表示如下:

$$GR = \frac{g}{dis + q} \quad (7)$$

利用上文参数数值设置的取值范围, 由式(7)计算得出的 GR 的值被映射到 $[0, 1]$ 上. 相对较优的

资源量化后的值接近 1, 而相对较差的资源量化后的值接近 0. 式(7)包括了资源的带宽、用户特征等因素. 如果侧重动态环境下的网络稳定性, 则应增大 W_s ; 如果重视网格系统的吞吐量, 应该增大 W_p ; 如果侧重网格任务的传输时间, 应该增大 W_b . 在这些参数之间权衡重要性可以优化网格调度模型, 提高网络动态适用性.

4 TR-BS 算法描述

任务和资源之间合理的匹配是一个复杂的过程. 本文为了简化这个过程, 设计了一种新的调度策略 TR-BS 算法. 算法的思想是综合任务的两类 QoS 需求量化结果, 计算出任务的综合需求参数, 根据参数大小将任务按非升序依次排序; 然后, 根据资源的相关参数计算出资源的综合性能参数, 并根据参数大小将资源按非升序排列. 对任务队列中的任务在资源序列中找出对应 QoS 参数值相当的资源进行匹配. 在资源选择的过程中, 如果在所选定的范围内找不到合适资源, 则在这个范围的基础上, 将资源综合性能参数值 (GR) 增加 ϵ (ϵ 是一个很小的数) 的范围内查找. 如果查到资源的 GR 最大值还没有找到合适的资源, 则在任务 QoS 对应资源范围的基础上减少 ϵ 继续查找, 直到找到对应的资源为止. 如果所有的资源都被查找过也没有合适的资源, 该任务在本次匹配中将被放弃, 参与下一次的调度. 在调度的过程中, 第一次被调度的任务优先级为 0, 任务被重新调度一次, 优先级提高一个级别, 级别高的任务排在任务列的最前面. 这意味着任务在下一轮的调度中被首先考虑.

本算法的优点就是在匹配的过程中, 任务的 QoS 值决定了它所匹配的资源 QoS 水平, 即任务找到的资源和任务本身的 QoS 水平相当, 这样, 低 QoS 要求的任务不会占据具有高性能的资源, 导致高 QoS 要求的任务因为没有合适的资源而等待的情况. 由于本算法的思想是将两个综合参数值相近的任务和资源匹配, 这样可以很好地平衡资源之间的负载, 减少任务的完成时间.

TR-BS 算法的具体执行过程如下:

输入: 该次调度的任务集合 T_o 和可用资源集合 M_o .

输出: T_o 中任务在可用资源之上的一个匹配对.

步骤 1: 网格资源提供者向网格系统注册中心注册资源的 QoS 信息和其他相关信息;

步骤 2: 网格用户向网格系统注册中心注册任务的 QoS 需求、权重参数、截止时间以及其他相关信息;

步骤 3: 收集有效计算资源 M_v , 初始化资源状态, 即将任一资源的可用时间置为 0 值;

步骤 4: 根据式(4)计算出所有任务综合 QoS 评估值 $G-task$ 并按非升序排列; 收集上一次调度剩下的任务, 并根据任务的优先级将任务依次排列;

步骤 5: 根据式(6)计算出所有资源综合性能评估值 GR 并按非升序排列;

步骤 6: 如果在所选定的范围内找不到合适资源, 则在这个范围的基础上, 将资源综合性能参数值 (GR) 增加 ϵ (ϵ 是一个很小的数) 的范围内查找. 如果查找到资源的 GR 最大值还没有找到合适的资源, 则在任务 QoS 对应资源范围的基础上减少 ϵ 继续查找, 直到找到对应的资源为止. 如果所有的资源都被查找过也没有合适的资源, 该任务在本次匹配中将被放弃, 参与下一次的调度.

步骤 7: 更新网格服务信息注册中心的数据, 以保持网格服务信息的同步更新.

5 实验仿真与结果分析

文中设计了 TR-BS 算法的仿真实验, 以验证该算法的正确性和有效性. 本实验编写了一个模拟程序, 仿真过程中需要的数据由程序自动生成. 为了体现该算法的有效性, 将该算法与 QoS 指导的 min-min 算法进行了比较. 将任务的 QoS 个数设为 10 个, $\epsilon=0.01$. 实验结果是程序运行 10 次的平均值.

图 1 是资源个数为 15、任务个数为 50~100 时的任务完成时间跨度. TR-BS 算法的时间跨度比 QoS 指导的 min-min 算法的执行时间少 20.61%. 因此, 相对 QoS 指导的 min-min 算法, TR-BS 算法能通过合理地调度, 将任务分配到与其 QoS 性能相当的资源上执行, 这种相当水平任务和资源的匹配减少了任务的时间跨度.

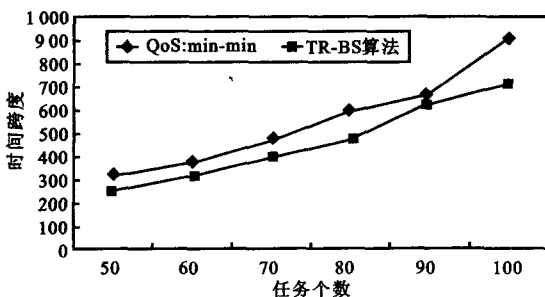


图1 时间跨度比较图表

图 2 显示了资源个数为 15、任务个数为 50~100 时的任务丢弃情况. 统计得, TR-BS 算法的任务丢弃率比 QoS 指导的 min-min 算法少 22.69%.

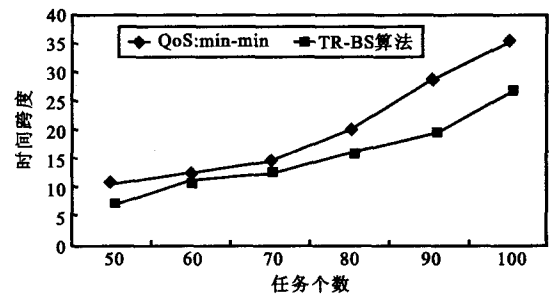


图2 任务丢弃情况

从图 3 中的曲线波动情况来看, TR-BS 算法资源负载曲线波动幅度较 QoS 指导的 min-min 算法小, 这种现象说明 TR-BS 算法能够平衡资源之间的负载, 不会出现很多任务集中在个别资源上导致资源负载失衡的情况.

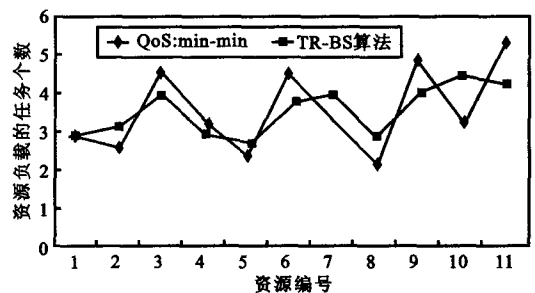


图3 资源负载情况

6 结束语

文中结合网格系统分布、异构和动态的网格环境特点, 考虑任务服务质量要求和资源综合性能的条件下, 针对如何在复杂的环境下快速解决任务调度的问题, 提出了一种新算法 TR-BS 算法. 该算法能够更好地协调任务和资源之间的匹配. 仿真实验证明 TR-BS 算法能够适应复杂的网格环境, 有效地为用户分配网格资源, 减少任务的时间跨度, 提高网格系统的吞吐量.

参考文献:

- [1] Foster I, Kesselman C, Tuecke S. The anatomy of the grid: enabling scalable virtual organizations[J]. International J. Supercomputer Application, 2001, 15(3): 200-222.
- [2] He X, Sun, Glaszewski. A QoS guided scheduling algorithm for grid computing[J]. Journal of Computer Science and Technology, 2003, 18(4): 442-451.

$3\text{M}^3/\text{H}$,且精度为1%,那么静态瞬时流量 $\delta_1 \leq 30\text{L}/\text{H}$.

图6为阈值比较法测得的静态瞬时流量,静态瞬时流量波动很大,而且存在测量错误点,远远不能满足分界流量为 $3\text{M}^3/\text{H}$,精度为1%的要求;图7为过零阈值模型测得的静态瞬时流量,基本上可以满足分界流量为 $3\text{M}^3/\text{H}$,精度为1%的要求,但是存在个别误差点;图8为过零阈值平滑模型测得的静态瞬时流量,完全可以满足始动流量为 $3\text{M}^3/\text{H}$,精度为1%的要求.

为了对比上述3种方法之间的稳定性和重复性^[11],文中对实测数据分别计算其期望、方差及标准差进行对比如表1所示.

表1 实测静态瞬时流量的数学统计

	期望	方差	标准差
阈值比较法	22.6134	128200	358.0520
过零阈值模型	1.4298	429.727	20.7299
过零阈值平滑模型	-0.0007	162.3995	12.7436

经过对3种方法的实际测量数据的数学统计分析可以看出,零阈值平滑模型实测数据的期望、方差和标准差要小,该系统的重复性和稳定性比较好,测量的数据更接近实际测量真实值,波动性很小,可以实现抑制静态瞬时流量波动,可以很好地解决静态漂移问题.

6 结束语

时差式超声流量的阈值比较模型,抗环境噪声和元器件老化的能力差,不能准确地判断接收超声信号到达的时刻,会产生静态漂移,一直制约时差式超声流量计精度的提高.文中对传统阈值比较法进行硬件改进和从系统角度对测得的数据进行软件优化,可以很好判断接收的超声信号到达的时刻点,较

准确测定顺流和逆流时间,有效抑制了静态漂移问题,提高时差式超声流量计的计量精度.

参考文献:

- [1] 蔡武昌.超声流量计的市场与应用[J].世界仪表与自动化,2007,10(11):14-15.
- [2] 罗再扬,李秋莲.浅谈超声流量计在原油记录中的应用[J].油气田地面工程,2007,26(4):36-37.
- [3] 兰纯纯.时差法超声波流量计的研究[D].重庆:重庆大学,2006.
- [4] 沈志辉,林伟军,张澄宇,等.用于井下超声石油流量计的互相关方法[J].声学技术,2008,27(6):845-849.
- [5] 李广峰.智能化超声波水流量计的硬件设计与制作[D].西安:西安理工大学,2000.
- [6] 蔡武昌,孙淮清,纪纲.流量测量方法和仪表的选用[M].北京:化学工业出版社,2004.
- [7] 刘存,黄建军.时差法超声流量计的几点改进[J].沈阳工业大学学报,2002,24(2):113-115.
- [8] 张涛,蒲诚,赵宇洋.传播时间法超声流量计信号处理技术进展述评[J].化工自动化及仪表,2009,36(4):1-7.
- [9] 姜勇.时差法超声波流量计设计与研发[D].杭州:浙江大学,2006.
- [10] 邱天爽,张旭秀,李小兵,等.统计信号处理-非高斯信号处理及其应用[M].北京:电子工业出版社,2004.
- [11] 张永瑞,刘振起,杨林耀,等.电子测量技术基础[M].西安:西安电子科技大学出版社,2007.

作者简介:

段允 男,(1986-),硕士研究生.研究方向为嵌入式应用.

王让定 男,(1962-),博士,教授,博士生导师.研究方向为信息安全、嵌入式应用.

朱莹 女,(1978-),硕士.研究方向为嵌入式应用、超声波应用.

姚灵 男,(1953-),教授级高工.研究方向为水计量技术.

(上接第204页)

- [3] Mohammad L, Analoui M. Grid - JQA: a QoS guided scheduling algorithm for grid computing[C]//Sixth International Symposium on Parallel and Distributed Computing (ISPDC'07). Austria,2007.
- [4] Kenli Li, Yan He, Xiaoling Liu. Security - driven scheduling algorithm based on eigen trust in grid[C]// Sixth International Symposium on Parallel and Distributed Computing. London,2005.
- [5] 李金忠,肖辉明. 网络服务 QoS 动态评估方法研究[J]. 井冈山学报, 2009, 30(2):50-53.

- [6] 吴雄奇,曾文华. 网络资源状态预测算法[J]. 计算机工程,2007,33(24):112-114.

作者简介:

韩云 男,(1985-),硕士研究生.研究方向为网络计算与云计算.

于炯 男,(1964-),教授,硕士生导师.研究方向为网络安全、网络与分布式计算.

张伟 男,(1984-),研究实习员.

王命全 男,(1986-),硕士研究生.