

## 基于商空间的层次式数据网格资源调度算法

夏纯中<sup>1,2</sup>, 宋顺林<sup>1</sup>

(1. 江苏大学 计算机科学与通信工程学院, 江苏 镇江 212013; 2. 江苏大学 信息化中心, 江苏 镇江 212013)

**摘要:** 为了解决传统数据网格调度算法在对层次式数据网格调度过程中出现的极易陷入局部最优值和收敛速度过慢的问题, 将粒计算的思想引入到网格调度中, 提出了一种基于商空间的层次式数据网格资源调度 QSHDGRA (quotient space theory based hierarchical data grid resource allocation) 算法。首先分析了层次式数据网格的特点, 接着提出一种基于业务请求平均等待时间和网络与节点资源利用均衡度的调和函数的调度问题模型, 随后设计了基于商空间的层次式最优资源调度算法。该算法的特点是在不同粒度上由粗至细地对网格业务进行调度, 从而保证不同业务的 QoS, 并实现系统全局最优资源分配。仿真实验表明, 算法可以显著地提升系统整体的吞吐率, 具有更快的收敛速度, 并具备线性扩展能力。

**关键词:** 数据网格; 资源调度; 分布式系统; 商空间; 粒子群算法

中图分类号: TP393.01

文献标识码: B

文章编号: 1000-436X(2013)06-0146-10

## Hierarchical data grid resource allocation based on quotient space theory

XIA Chun-zhong<sup>1,2</sup>, SONG Shun-lin<sup>1</sup>

(1. College of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China;

2. Information Center, Jiangsu University, Zhenjiang 212013, China)

**Abstract:** In order to solve the problems of falling into local optimum value and converging too slowly when allocating resources in hierarchical data grid using traditional algorithms, the granular computing was introduced and a quotient space theory based hierarchical data grid resource allocation (QSHDGRA) algorithm was proposed. Firstly, the characteristics of hierarchical data grid were analyzed. Secondly, a reconciling model of minimum average waiting time and maximum network and node resource utilization was defined, and then the QSHDGRA algorithm was designed. The algorithm can allocate resources from coarse granularities to fine ones, so it can guarantee the QoS of different businesses and make global optimal resource allocation. Simulation results show that QSHDGRA can improve overall system throughput with faster convergence speed and linear scalability.

**Key words:** data grid; resource allocation; distributed system; quotient space theory; particle swarm algorithm

### 1 引言

近年来, 数据网格凭借其强大的扩展能力被用于构建企业级数据库云平台。例如, 大型医疗集团信息集成平台<sup>[1]</sup>利用数据网格对集团内各医疗机构分布异构的医疗信息数据库进行集成和共享; 红帽

发布的 JBoss Enterprise Data Grid<sup>[2]</sup>更是将数据网格作为其企业云计算战略的重要基础构件。层次式数据网格是一种常见的数据网格架构, 由于大规模分布式系统具有小世界特性, 用户在对数据的使用上呈现出社团性和层次性, 即特定区域内的用户只对特定部分的数据最感兴趣, 而层次式数据网格可以

收稿日期: 2012-09-13; 修回日期: 2013-04-25

基金项目: 十一五国家科技支撑计划基金资助项目(2010BAI88B00); 国家自然科学基金资助项目(61272074); 江苏省自然科学基金资助项目(BK2011464); 江苏省普通高校研究生科研创新计划基金资助项目(CXLX11\_0591)

**Foundation Items:** The Eleventh Five-Year National Science and Technology Support Program (2010BAI88B00); The National Natural Science Foundation of China (61272074); The Natural Science Foundation of Jiangsu Province (BK2011464); Jiangsu College Graduate Research and Innovation Plan (CXLX11\_0591)

很好满足这一需求。

在企业应用领域，用户对数据网格的服务质量（QoS, quality of service）有着和科学计算截然不同的需求。网络连接带宽和不同网格节点处理能力的差异导致数据网格无法像传统的企业数据中心那样为业务系统提供实时、可靠的数据服务。层次式数据网格资源调度的目的就是为网格中各节点和不同级别业务分配带宽，保障各级业务的服务质量，并使系统整体资源利用率最优。

层次式数据网格资源调度可以采用现有的数据网格调度算法，常见的数据网格调度算法大致分为 4 类。1) 基于综合指标直接选择法。Cheng<sup>[3]</sup>提出一种基于 Min-Max 的负载均衡算法，将文件和网络带宽分为 Min 和 Max 2 类，调度时根据文件所属类别到相应站点下载。Mistarihi 等人<sup>[4]</sup>使用响应时间、可靠性和安全构成一个评价指标，并使用层次分析法确定各指标的权重。Du 等人<sup>[5]</sup>提出了一种基于可靠度的指标来减少任务完成时间和执行代价。Qu 等人<sup>[6]</sup>提出网络负载和服务容忍度 2 个目标函数，使用权重因子平衡 2 个目标函数。Tang 等人<sup>[7]</sup>提出一种副本价格模型，模型综合考虑了 CPU、内存、网络的延时、带宽和可靠性等多种因素，并通过拍卖协议选取传输代价最小的副本。2) 基于机器学习法。文献[8,9]都采用神经网络构建数据传输时间预测模型，使用日志数据中不同大小文件在不同网络带宽下的传输时间作为样本对模型进行训练。文献[10]采用  $k$ -NN 算法对不同节点磁盘 IO 和网络延时下的文件下载时间进行学习。文献[11]中，Almuttairi 等人提出基于  $K$  均值聚类对站点进行分类，采用基于灰度的粗糙集理论解决安全、可靠性、成本和响应时间等属性不确定的问题。3) 基于时间序列法。Li 等人<sup>[12]</sup>使用灰色理论预测副本响应时间，并使用马尔科夫链预测副本的可靠性。Wu 等人<sup>[13]</sup>提出一种状态模糊评估策略，并使用灰色  $GM(1,1)$ 模型动态地预测服务时间。4) 基于进化计算法。Munoz<sup>[14]</sup>提出基于粒子群优化算法，根据网络延迟和带宽确定传输代价，使用访问率和传输代价作为最优化评价函数。Xiong<sup>[15]</sup>提出了副本选择算法的 QoS 模型，采用层次分析法分析不同 QoS 指标的重要程度，并以此作为不同指标的权重向量。提出了一种基于 MapReduce 的并行遗传算法加速运算。算法主要考虑 3 个 QoS 指标，即平均传输率、平均带宽和可靠性。直接使用上述算法对层次

式数据网格进行调度时，由于算法中所有节点均在同一层面，当网格节点数量很大时，极易出现以下 2 个问题：①算法难以保证系统全局最优，这主要是由于算法缺少在不同粒度、不同层面上对网格流量进行调度，没有充分利用层内和层间节点对不同优先级的任务进行优化。②数据网格资源调度问题属于 NP 问题，当网格内节点的数量较多时，全局资源调度算法收敛过慢，难以满足企业应用对数据实时性的要求。

粒计算是一种通过对复杂问题进行粒化分解，从而降低问题复杂度的方法论，它是模拟人类在解决复杂问题时能够由粗至细，多层次地观察和分析问题的能力。张铃和张钺提出使用商空间理论对粒度空间进行描述。商空间理论通过保真原理和保假原理在论域元素上构造一系列不同粒度的商空间，形成一个分层递阶的商空间链，从而简化原始问题的求解<sup>[16]</sup>。

本文首次引入粒计算的思想解决传统数据网格资源调度算法在调度层次式数据网格时存在的问题，提出了一种基于商空间的层次式数据网格资源调度 QSHDGRA (quotient space theory based hierarchical data grid resource allocation) 算法。其主要创新在于：采用商空间粒度计算的方法在不同粒度层面对数据网格系统流量进行调度，保障不同优先级业务的 QoS 的需求，此外调度算法还兼顾了数据网格整体节点负载和网络流量的均衡，实现了系统资源利用的全局最优化。

## 2 层次式数据网格模型

### 2.1 体系结构

层次式数据组织是一种常见的数据管理方式。如图 1 所示，在层次式数据网格中，所有节点组成一个层次化的树状网络，根节点拥有全部数据副本，根节点将所有数据副本平均分配到各个二级节点，二级节点再将数据副本平均分配到下属的三级节点。一级节点定期监控各节点和各条链路的负载情况，采用特定算法分配各个节点的处理带宽的能力，实现系统访问性能的全局最优化。

为了满足企业应用集成领域中不同业务 QoS 的要求，将数据网格的业务请求分为 3 种 QoS 级别：第一类业务优先级最高，要求能够立即响应；第二类业务优先级为中，允许在一定时间内完成响应；第三类业务优先级最低，当系统繁忙时，可以暂停响应。

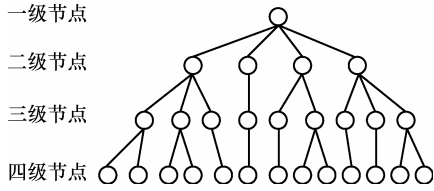


图 1 层次式数据网格节点组织

层次式数据网格有以下 3 个优点。1) 网格数据具有高可靠性。根节点和各级节点的数据互为异地备份。每一个数据均有 3 个副本，且副本互为异地备份。2) 网格系统具有很好的容错性。任何节点的单点故障都不会影响系统运行。例如，当根节点发生故障时，其他副本节点可继续为用户提供数据访问服务。3) 网格系统具有负载均衡能力。根节点和各级节点构成一个分布式虚拟集群。当某个节点或某条路径访问负载过高时，系统会自动地将访问请求分配到其他节点上。

### 2.2 问题形式化定义

在层次式数据网格有向图  $G(V,E)$  中，采用  $M/M/1$  排队模型定义系统资源调度问题。

**定义 1** (节点请求平均等待时间)。单位时间内到达节点  $j$  的请求数为  $\lambda_j$ ，节点  $j$  的处理能力为  $\mu_j$ ，则节点  $j$  的平均请求等待时间为  $T_j$  为

$$T_j(\lambda_j) = \frac{1}{\mu_j - \lambda_j} \quad (1)$$

**定义 2** (节点资源利用率)。设  $S_j$  为单位时间内到达节点  $j$  的请求数占其最大处理能力的比例，则该节点资源利用率为

$$S_j = \frac{\lambda_j}{\mu_j} \quad (2)$$

**定义 3** (网络资源利用率)。设从节点  $i$  出发经过路径  $p_{ij}$  到达节点  $j$  的请求数为  $\lambda_{ij}$ ，从节点  $j$  出发经过路径  $p_{ji}$  到达节点  $i$  的请求数为  $\lambda_{ji}$ ，路径  $p_{ij}$  的最大吞吐率为  $l_{ij}$ ，则路径  $p_{ij}$  的网络资源利用率为

$$L_{ij} = \frac{\lambda_{ij} + \lambda_{ji}}{l_{ij}} \quad (3)$$

**定义 4** (请求流量矩阵)。用请求流量矩阵  $(\lambda_{QoS m, ij})_{N \times N} (m=1,2,3)$  表示系统内从节点  $i$  出发到达节点  $j$  的不同 QoS 级别流量的分配情况。3 种 QoS 级别流量矩阵相加，即为节点  $i$  到  $j$  的流量。

$$(\lambda_{QoS m, ij})_{N \times N} = (\lambda_{QoS 1, ij})_{N \times N} + (\lambda_{QoS 2, ij})_{N \times N} + (\lambda_{QoS 3, ij})_{N \times N} \quad (4)$$

**定义 5** (系统资源调度的多目标优化问题)。

层次式数据网格系统资源调度问题可以分为以下 3 个子问题：一是为不同 QoS 级别用户分配带宽，使用户平均请求等待时间最小化；二是节点负载利用率最大化。即充分利用冗余节点实现请求负载均衡，避免请求过分集中到某些节点上；三是网络流量利用率最大化。即充分利用空余网络链路，避免造成某些路径流量过大，而另外一些路径却未充分使用。

综上所述，系统资源调度的全局最优化目标为确定请求流量矩阵  $(\lambda_{QoS m, ij})_{N \times N}$  使得不同 QoS 级别业务的平均请求等待时间最小化，同时实现节点资源利用率和网络资源利用最大化。

$$\min T(\lambda_{QoS m}) = \frac{1}{\lambda_{QoS m}} \sum_{j=1}^N \lambda_{QoS m, j} \cdot T_j(\lambda_{QoS m, j}) \quad (5)$$

$$\max HS(\lambda) = - \sum_{j=1}^N S_j \ln S_j \quad (6)$$

$$\max HL(\lambda) = - \sum_{k=1}^M L_k \ln L_k \quad (7)$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^N \lambda_{QoS m, ij} = \lambda_{QoS m, j}, m = 1, 2, 3 \\ \sum_{m=1}^3 \lambda_{QoS m, j} = \lambda_j \\ \sum_{j=1}^N \lambda_j = \lambda \\ S_j < 1, L_k < 1 \\ \lambda_{QoS m, ij} \geq 0, \lambda_i \geq 0, \lambda_j \geq 0, \lambda_{ij} \geq 0 \\ \mu_i \geq 0, l_{ij} \geq 0 \end{cases} \quad (8)$$

式(6)和式(7)使用信息熵衡量网格中节点负载的均衡度和路径流量的均衡度。第一个约束条件表示系统请求流量守恒，第二个约束条件表示节点负载分配后的流量守恒，第三个约束条件保证队列平均长度不会无限增大。

**定义 6** (全局最优目标调和函数)。采用线性加权和法将定义 5 中的多目标问题转换为单目标问题，定义层次式数据网格系统全局最优资源调度目标函数为

$$\min Z(\lambda) = \alpha T(\lambda_{QoS 1}) + \beta T(\lambda_{QoS 2}) + \gamma T(\lambda_{QoS 3}) - \frac{1}{\delta} HL(\lambda) - \frac{1}{\varepsilon} HS(\lambda) \quad (9)$$

式(9)是系统中各 QoS 用户请求的平均等待时间和网络与节点资源利用均衡度的调和函数。

### 3 基于商空间的资源调度算法

#### 3.1 传统优化算法及其弊端

本文的最优化问题是一个非线性规划问题，且属于 NP 问题，采用启发式计算的方法可以快速获得最优解。粒子群算法是一种常用的基于进化计算的启发式算法，其主要特点是个体数目少，计算简单，顽健性好和可并行计算。使用粒子群算法首先要确定粒子的编码方式和适应度函数。本文以数据网络节点间的有向连接数作为解空间维度，每一维对应一条有向连接，其取值为分配给该连接请求流量。每个粒子对应的  $D$  维向量  $X$  表示系统资源调度问题的一个解，适应度函数为式(9)所定义的目标函数。

粒子群算法在求解本文最优化问题中暴露出以下的弊端：由于问题的解空间的向量维度太大，使得算法收敛到全局最优值的速度非常缓慢，并且极易陷入局部最优值。

#### 3.2 基于商空间的优化算法

在层次化网络中，随着网络节点数目的增多，节点间的连接数量呈几何级数增长，网络拓扑结构越来越复杂，资源调度算法的搜索空间也越来越大。本文的主要思想就是使用粒度计算的思想来简化求解层次化网络最优资源调度问题的复杂度。通过对原始最优化问题进行层层分解，在不同粒度、不同层面上对问题进行优化，以求加快求解速度，并获得全局最优值。

##### 3.2.1 构建商空间

给定一个  $L$  层的层次网络的网络图  $G(V,E)$ ，用  $l(v)$  表示节点所属的层次。根节点属于第一层，即  $l_1=1$ ，其余各层次关系为  $l_1 < \dots < l_k \dots < l_L$ 。用  $p(v_i, v_j)$  表示  $v_j$  是  $v_i$  的子节点。下面基于等价关系商空间理论，构建等价类对节点进行划分，构成商空间。

**定义 7** 等价关系  $R(l_k)$ 。

$$R(l_k) = \left\{ \langle x, y \rangle \left| \begin{array}{l} x \in V, y \in V, \exists z \in V, l(x) \geq l_k, l(y) \geq l_k, \\ l(z) = l_k, p(z, x), p(z, y) \end{array} \right. \right\}$$

该等价关系的实质是将某一级节点及其各自所有子节点看作一个整体，即一个粒子，和上级节点归为同一等价类。当  $l_i$  从 1 增大到  $k$  时，可以在各个级别层次上构建不同粒度的网络拓扑，形成分层递阶的商空间链。

定义等价关系  $R(l_k)$  对应的商空间为  $V_k = \{v_1^k,$

$v_2^k, \dots, v_n^k\}$ ,  $k=1,2,\dots,L$ 。  $V_1$  表示最粗粒度的空间，即将所有节点集合看作一个粒子  $v_1^1$ ，  $V_L$  表示最细粒度的空间，其中节点集合  $V_L = \{v_1^L, v_2^L, \dots, v_n^L\}$  和原始网络中的节点一一对应。按照节点集合由粗到细可得到  $V_1 > V_2 > \dots > V_L$  这样一个分层递阶的商集。

在  $V_k$  上定义边集合  $E_k$ ，  $(v_i^k, v_j^k) \in E_k \Leftrightarrow \exists v_i^L \in v_i^k, v_j^L \in v_j^k, e(v_i^L, v_j^L)$ 。该定义表示在层次  $k$  的粒度上节点间存在连接的等价条件为这对粒子包含的子节点在层次  $L$  的原始网络中存在连接。

记数据网络拓扑商空间为  $G_k(V_k, E_k)$ ,  $k=1,2,\dots,L$ 。根据网络粒度进行排序，  $G_1 > G_2 > \dots > G_L$  构成一个分层递阶商空间链。

##### 3.2.2 资源调度算法

在构建了商空间之后，就可以由粗到细地逐层对系统流量进行分配。下面对不同粒度节点单位时间请求率  $\lambda$  和请求处理率  $\mu$  进行定义。对于位于第  $k$  层的某节点  $v_i^k$ ，其单位时间请求率  $\lambda_i^k$  为其所有子节点的请求率之和，其请求处理率  $\mu_i^k$  为其所有子节点的请求处理率之和与所有子节点链路的最大通行能力之和的最小值。在商空间  $G_k$  中，最优分配算法对应的解向量表示为  $X_k$ 。

基于商空间的层次式数据网络资源调度算法形式化定义如下。

**问题：**给定层次式数据网络  $G(V,E)$ ，确定其上的各优先级业务的最优请求流量矩阵  $(\lambda_{QoS_{m,ij}})_{N \times N}$ ，使得系统全局资源调度目标值  $Z(\lambda)$  最小。

算法步骤描述如下。

1) 构建商空间。输入原始层次化网络模型  $G$ ，给出网络节点数量  $V$ ，网络拓扑结构  $E$ ，各连接链路最大负载  $L$  和节点的最大负载  $S$ 。根据网络层次结构确定等价关系  $R$ ，并构建分层递阶商空间链  $G_1 > G_2 > \dots > G_L$ 。

2) 确定求解最优化算法的粒子群规模  $m$  和算法参数，包括惯性因子  $w$ 、学习因子  $c$  和速度限值  $S$ 。设定最大迭代次数  $n$  和迭代终止阈值  $\varepsilon$ 。

3) 由于  $G_1$  是将全体节点看作一个整体，没有进行划分，因此从第二层商空间开始进行递归求解。设  $k=2$ ，转到 4)。

4) 在商空间  $G_k$  中，根据节点数量  $V_k$  和连接数量  $E_k$  构建流量矩阵作  $\lambda_{QoS_{m,ij}}^k$  为解向量  $X_k$ 。使用粒子群算法完成一次搜索，求得本层粒子的一个最优

解。查看该空间是否有更细粒度的空间，如果有转到 5)，否则转到 6)。

5) 进入更细一层商空间  $G_{k+1}$ ，构建流量矩阵  $\lambda_{QoS_m}^{k+1}$ ，由于上一层空间的搜索有陷入局部最优值的可能，因此本层使用 2 个粒子群按照不同的策略进行搜索。第一个粒子群中增加上一层空间求得的最优解作为本次求解的约束条件，即由上一层粒子分解而得各粒子负载之和不能超出上一层粒子获得的最优分配值。第二个粒子群则不增加约束条件进行搜索。当 2 个粒子群分别完成最优化搜索之后，比较其最优解。如果不加约束条件的粒子群获得的解更优，则将该解合成为上一层空间的最优解，重复步骤 5)，跳转到上一层空间重新搜索。如果加了约束条件的粒子群获得的解更优，则该最优解为本层粒子的最优解。查看本层是否有更细粒度的空间，如果有继续 5)，跳转到下一层空间进行搜索，否则转到 6)。

6) 此时已完成最细一层商空间，得到原始网络空间的资源分配流量矩阵  $\lambda_{QoS_m}$ 。系统根据最优值定义的请求流量矩阵调度各级 QoS 业务请求流量。

### 3.2.3 算法实例

下面结合一个例子来阐述算法过程，给出如图 2 所示的层次化网络。网络共有 15 个节点，为了便于描述，网络采用树状结构。

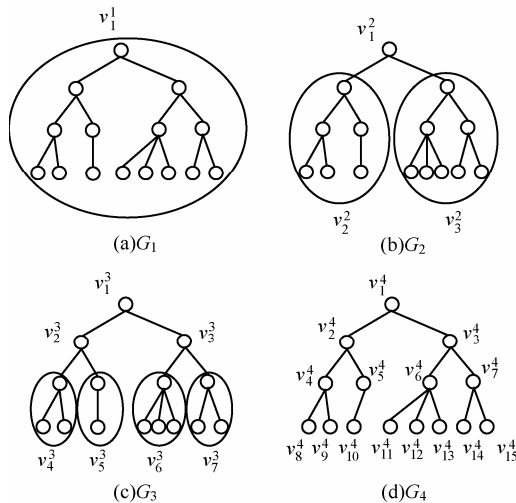


图 2 商空间构建过程

原始网络中所有节点全体  $G_1$  作为初始商空间，节点集记为  $V_1$ 。

$$V_1 = \{v_1^1 = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15)\}$$

首先使用等价类  $R(2)$  对原始空间进行划分，得

到商空间  $G_2$ ，其商集  $V_2$  如下。

$$V_2 = \left\{ \begin{aligned} v_1^2 &= (1), v_2^2 = (2, 4, 5, 8, 9, 10), \\ v_3^2 &= (3, 6, 7, 11, 12, 13, 14, 15) \end{aligned} \right\}$$

这样原始空间划分成包含 3 个元素的粗粒度网络图，其边集合为

$$E_2 = \{e(v_1^2, v_2^2), e(v_1^2, v_3^2), e(v_2^2, v_3^2)\}$$

$(V_2, E_2)$  构成商空间  $G_2$ ，其拓扑结构如图 3 所示。

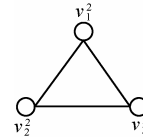


图 3  $G_2$  商空间拓扑结构

其流量分配矩阵为

$$\lambda^2 = \begin{pmatrix} \lambda_{11}^2 & \lambda_{12}^2 & \lambda_{13}^2 \\ \lambda_{21}^2 & \lambda_{22}^2 & \lambda_{23}^2 \\ \lambda_{31}^2 & \lambda_{32}^2 & \lambda_{33}^2 \end{pmatrix}$$

矩阵中每个元素为基本粒中的连接流量之和，例如  $\lambda_{2,1}^2 = \lambda_{2,1} + \lambda_{4,1} + \lambda_{5,1} + \lambda_{8,1} + \lambda_{9,1} + \lambda_{10,1} + \lambda_{11,1}$ 。确定好  $G_2$  空间的解向量后，根据粒子群算法求得该层空间的最优资源分配解。

接下来，使用等价类  $R(3)$  对  $G_2$  进行划分，得到商空间  $G_3$ ，其商集  $V_3$  如下。

$$V_3 = \left\{ \begin{aligned} v_1^3 &= (1), v_2^3 = (2), v_3^3 = (3), v_4^3 = (4, 8, 9), v_5^3 = (5, 10), \\ v_6^3 &= (6, 11, 12, 13), v_7^3 = (7, 14, 15) \end{aligned} \right\}$$

其流量分配矩阵为

$$\lambda^3 = \begin{pmatrix} \lambda_{11}^3 & \lambda_{12}^3 & \lambda_{13}^3 & \lambda_{14}^3 \\ \lambda_{21}^3 & \lambda_{22}^3 & \lambda_{23}^3 & \lambda_{24}^3 \\ \lambda_{31}^3 & \lambda_{32}^3 & \lambda_{33}^3 & \lambda_{34}^3 \\ \lambda_{41}^3 & \lambda_{42}^3 & \lambda_{43}^3 & \lambda_{44}^3 \end{pmatrix}$$

同样，使用 2 个粒子群算法求得  $G_3$  空间中的最优解，第一个粒子群需要满足  $G_2$  空间已经求得的最优解的约束条件，即  $\lambda_{21}^3 + \lambda_{41}^3 + \lambda_{51}^3 \leq \lambda_{21}^2$ 。

最后，使用等价类  $R(4)$  对  $G_3$  进行划分，得到最细粒度的商空间  $G_4$ ，其商集  $V_4$  如下

$$V_4 = \left\{ \begin{aligned} v_1^4 &= (1), v_2^4 = (2), v_3^4 = (3), v_4^4 = (4), v_5^4 = (5), v_6^4 = (6), \\ v_7^4 &= (7), v_8^4 = (8), v_9^4 = (9), v_{10}^4 = (10), v_{11}^4 = (11), \\ v_{12}^4 &= (12), v_{13}^4 = (13), v_{14}^4 = (14), v_{15}^4 = (15) \end{aligned} \right\}$$

由于该空间等同于原始问题空间，其求得最优

解就是原问题的最优解。

### 3.3 算法分析

#### 3.3.1 命题证明

**命题 1** 对于层次式数据网格采用基于等价关系构建的分层递阶商空间链，可以求得原始问题的最优解。

**证明** 对于原始网络，商集 $[X]$ 中的元素就是各个节点，商结构 $[T]$ 是原始网络拓扑，假设商空间链有 $L$ 层。原始问题存在最优解，即在分层递阶商空间链中，最细粒度的空间 $([X_L], [T_L])$ 存在最优解。采用等价关系 $R(L-1)$ 对其进行划分，得到上一层商空间 $([X_{L-1}], [T_{L-1}])$ ，根据商空间的保真原理可知该空间中必包含原始问题的最优解。反复应用商空间保真原理可得最粗粒度商空间 $([X_2], [T_2])$ 存在最优解。由此可得原始空间最优解必包含在各级空间的最优解中，因此基于分层递阶商空间链逐层搜索可以求得原始问题的最优解。

**命题 2** 在分层递阶商空间中求解的过程，是一个逐步逼近最优解的过程。

**证明** 采用反证法。假设在某一粒度的空间 $([X_L], [T_L])$ 中，求得一个当前最优解，在该粒度空间最优解的约束下，更细粒度的下级空间 $([X_{L+1}], [T_{L+1}])$ 可以得到一个更优解。假设该解不是最优解，说明超出上级空间约束还有最优解，将该最优解的分量相加，由保真原理得到上级空间解就比之前求得的最优解更优，与之前的解是最优解的假设矛盾。

由此命题可以得出，如果要求出全局最优解，那么每一层的解都必须是当前层内的全局最优解。故此在算法中采用了双粒子群，一个粒子群基于约束条件搜索，另一个粒子群不受约束搜索，这样可以充分保证解空间的完备性。

#### 3.3.2 收敛性分析

分析 QSHDGRA 算法收敛性需要用到 Solis 的随机优化算法以概率 1 收敛于全局最优解的充分条件。

**命题 3** 假设 QSHDGRA 求解的目标函数 $f$ 是可测函数，其解空间 $S$ 为可测集，那么，QSHDGRA 算法以概率 1 收敛于全局最优解。

**证明** 由 Solis 收敛定理知，只需证明 QSHDGRA 满足 Solis 收敛假设 1 和假设 2 即可。

1) QSHDGRA 算法中商空间由粗到细的搜索过程是一个对搜索空间逐步压缩的过程，每一层搜索

的最优值逐步逼近全局最优值。迭代函数 $f$ 可归结为

$$F(p_g(t), p_i(t)) = \begin{cases} p_g(t), f(p_g(t)) \leq f(x_i(t)) \\ x_i(t), f(p_g(t)) > f(x_i(t)) \end{cases} \quad (10)$$

其中， $t$  为进化代数，算法的解序列为 $\{p_g(t)\}_{t=0}^{\infty}$ ，可以保证其趋向最优值，因而容易证明其满足 Solis 收敛假设 1。

2) 设各个粒度网络中最优解的样本空间的并必包含 $S$ ，即 $S \subseteq \bigcup_{i=1}^L M_{i,t}$ 。其中 $M_{i,t}$ 为第 $t$ 代粒子 $i$ 的样本空间支撑集。QSHDGRA 算法在每一层空间采用 2 个粒子群进行搜索，第一个粒子群保证搜索空间的收敛，第二个粒子群将保证种群多样性，随机搜索空间而不进行收敛，设第一个粒子群支撑集的并集为 $\alpha$ ，第二个粒子群支撑集的并集为 $\beta$ 。由于第二个粒子群搜索的随机性，必然存在整数 $t_1$ ，使得当 $t > t_1$ 时， $\beta \supseteq S$ 。因此，对于 QSHDGRA 算法，存在整数 $t_2$ ，使得 $t > t_2$ 时 $\alpha \cup \beta \supseteq S$ 。定义 $S$ 的任意 Borel 子集 $A = M_{i,t}$ ，则有 $\nu(A) > 0$ ， $\mu_i[A] = \sum_{i=1}^L \mu_{i,t}[A] = 1$ ，即 $\prod_{t=0}^{\infty} (1 - \mu_t[A]) = 0$ 。所以 QSHDGRA 算法满足 Solis 收敛假设 2。

综合 1) 和 2)，由 Solis 收敛定理可得 QSHDGRA 算法是一个全局收敛算法，能以概率 1 全局收敛。

#### 3.3.3 复杂度分析

粒子群算法的收敛速度与其解空间规模相关，给定网络 $G(V, E)$ ，节点数量为 $n$ ，解空间维度为 $m = n(n-1)$ ，则收敛的时间复杂度为 $O(s^m)$ ，其中 $s$ 为解向量中每一维度的取值范围。基于等价关系商空间模型构建 $L$ 层分层递阶的商空间链 $G_1 > G_2 > \dots > G_L$ 。设在商空间 $G_k$ 中，解空间维度为 $m_k$ ，算法收敛的时间复杂度为 $O(s_k^{m_k})$ ，则整个算法的时间复杂度为 $\sum_{k=1}^L O(s_k^{m_k})$ 。在算法求解过程中，虽然解空间维度逐步增大，即 $m_1 < m_2 < \dots < m_L$ ，但由于搜索空间不断减小，每一维度的取值范围也逐步减小，即 $s_1 > s_2 > \dots > s_L$ ，因此最优情况下，算法能够以几何速度收敛。

## 4 仿真实验与分析

### 4.1 仿真环境与实验方法

仿真实验采用 GridSim 仿真软件。仿真网络拓扑如图 2 所示：一级节点，即根节点 1 个；二

级节点 2 个，采用 100 Mbit/s 专线和一级节点链接；三级节点 4 个，采用 50 Mbit/s 专线连接到所属的二级节点；四级节点 8 个，采用 10 Mbit/s 专线连接到所属的三级节点。各节点内部网络是 1 000 Mbit/s 网络，各节点之间均采用虚链路实现路由连通。仿真文件大小为 1~5 MB 之间的随机值，系统共产生 1 000 个文件，按层次关系分布到各个节点中。系统采样单位时间为 5 min，即 300 s，系统请求率为每单位时间生成的请求数，节点处理能力的单位为每单位时间处理的事务数，四级节点的最大处理能力分别为 300、200、120 和 80。

实验包含 4 个部分：首先分析了不同 QoS 参数对不同 QoS 服务吞吐率和响应时间的影响；其次分析了全局均衡参数对系统整体性能的影响；接着分析了算法在不同种群数量、迭代次数和网络规模下的收敛速度；最后将本文的 QSHDGRA 算法和 MinTime 最小时间算法、Random 算法的性能进行了横向比较。

### 4.2 不同 QoS 级别性能分析

首先比较不同 QoS 参数对不同级别业务的吞吐率和平均响应时间的影响。对式(9)中的均衡参数

$\delta$  和  $\varepsilon$  取零，分别考察  $\alpha:\beta:\gamma$  为 1:1:1, 5:3:2 和 6:3:1 3 种情况下性能差别。

从图 4 和图 5 中可以看出，随着系统请求率由 400 增大到 2 000，服务响应率也相应地由 400 增加到 1 370。当系统请求率小于 800 时，系统处于轻载，所有的请求基本都能够及时响应，不同 QoS 参数比下的 3 种业务的吞吐率大致相同，平均等待时间也大致相等。当系统请求率增大到 1 200 时，系统负载加大，不同 QoS 参数比下的 3 种业务吞吐率出现明显差异。当  $\alpha:\beta:\gamma=1:1:1$  时，3 种业务的吞吐率仍然保持大致相当，但当  $\alpha:\beta:\gamma=5:3:2$  和  $\alpha:\beta:\gamma=6:3:1$  时中，3 种业务的吞吐率依次减少，业务平均等待时间依次增加。当系统请求率大于 1 600 时，已经接近系统的满载负荷，第一种参数比下 3 种业务吞吐率仍然近似相等，而后 2 种参数比下的吞吐率大致比例为 5:3:2 和 6:3:1，且第二、三类业务的响应时间明显大于第一类业务。由此可见，QSHDGRA 算法通过调节参数  $\alpha:\beta:\gamma$  可以控制不同优先级业务的流量，从而保障高优先级业务的吞吐率和相应时间约束。

### 4.3 全局均衡性能分析

本节分析全局均衡参数对系统性能的影响，即

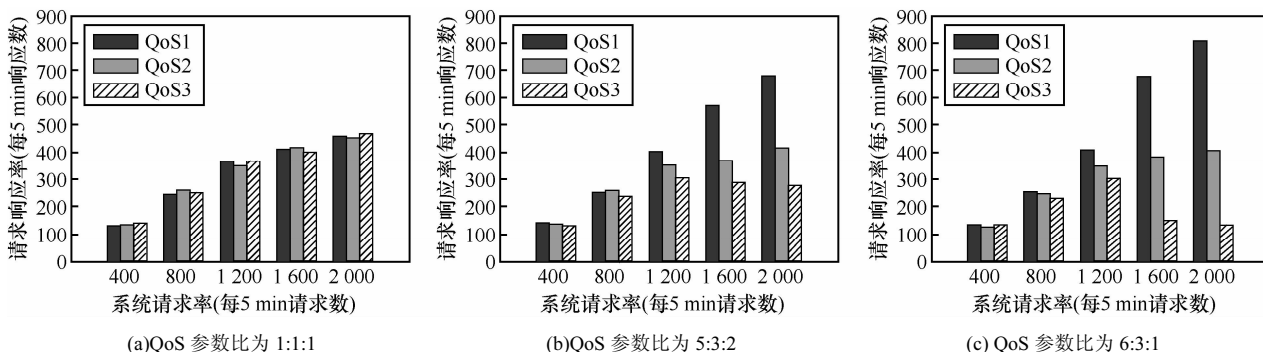


图 4 不同 QoS 参数比时各级别业务吞吐率比较

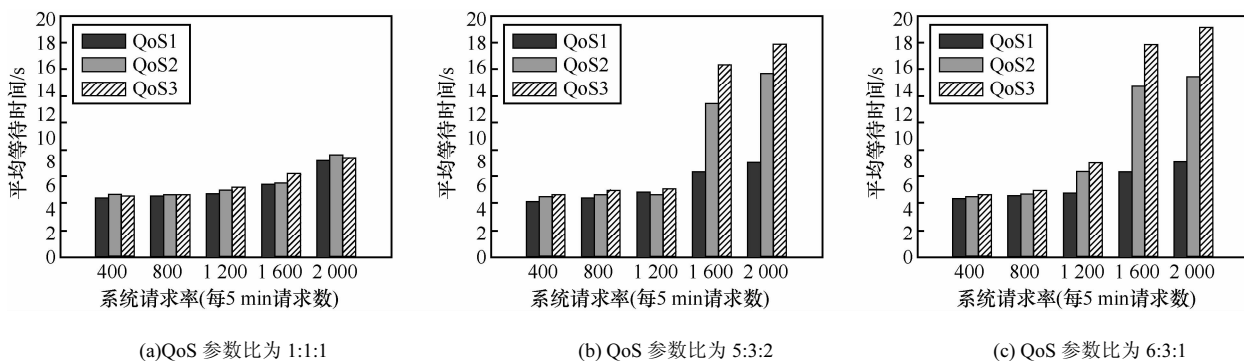
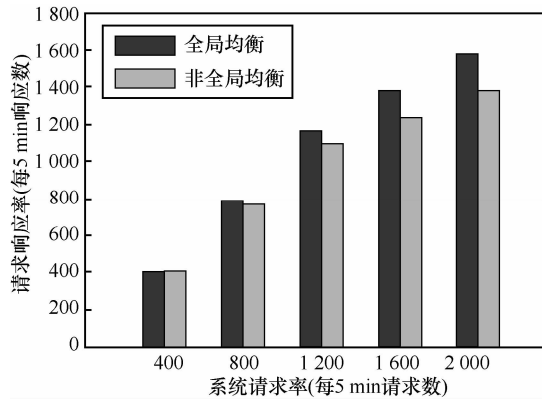


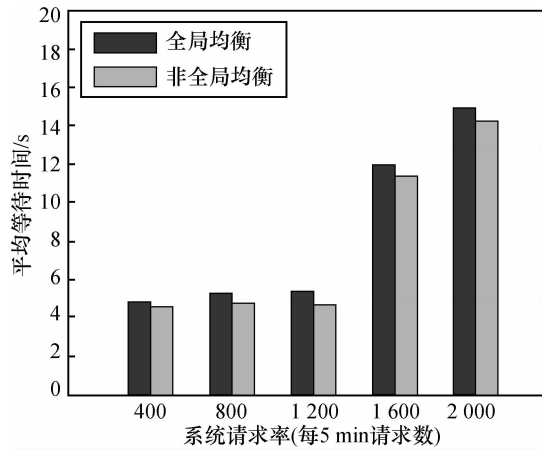
图 5 不同 QoS 参数比时各级别业务平均等待时间比较

式(9)中的均衡参数  $\delta$  和  $\epsilon$  对系统吞吐率和平均等待时间的影响，公式中  $\alpha:\beta:\gamma$  取固定值 5:3:2。

从图 6(a)中可以看出，随着系统请求率逐步增大，使用均衡参数比不使用均衡参数可以获得更高的响应率，即使在系统重载时，仍可以进一步提升响应率。从图 6(b)中可以看出，在系统轻载时，两者的平均响应时间大致相当，当系统重载时，使用均衡参数时平均响应时间会略大，这是由于平均响应时间在目标函数中的比例已经降低，系统向着全局吞吐率更大的目标进行优化。



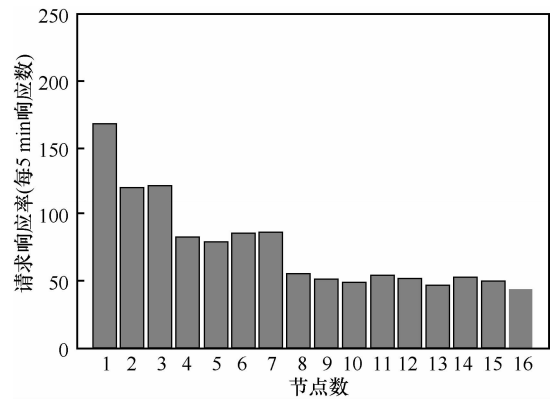
(a) 服务响应率比较



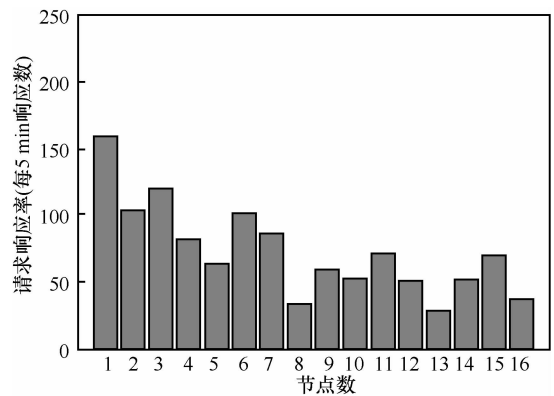
(b) 平均等待时间比较

图 6 使用和不使用均衡参数下吞吐率和平均等待时间的比较

图 7 显示了使用均衡参数和不使用均衡参数时各节点服务响应率的差别。此时系统请求率为 1 200，处于轻载状态。从图中可以看出，使用均衡参数和不使用均衡参数相比，各级节点的负载更加平均。由此可见，全局均衡参数可以很好地平衡各个节点的工作负载，有效避免某些节点处于繁忙状态，而另一些节点处于空闲状态，从而进一步提升系统整体的吞吐率。



(a) 使用均衡参数



(b) 不使用均衡参数

图 7 使用和不使用均衡参数下各节点服务响应率比较

### 4.4 算法收敛速度分析

本节分析 QSHDGRA 算法和不使用商空间的 PSO 算法在不同种群数量、迭代次数和网络规模对算法收敛速度的差别。

从图 8(a)中可以看出，粒子群的种群数量由 20 增大到 100 时，QSHDGRA 算法的收敛速度快于 PSO 算法，不过当种群数量大于 60 后，速度减少的幅度也相应减小。在图 8(b)中，当迭代次数由 200 增加到 1000 时，QSHDGRA 算法比 PSO 算法得到的适应值更优，且可以更快地收敛到最优值。综合图 8(a)和 8(b)可得，种群数量为 60，迭代次数为 600 时，系统可以在可接受的时间内获得近似的最优值。图 8(c)显示了网络规模扩大时，算法计算时间的变化。当网络规模由 15 个节点增大到 255 个节点时，QSHDGRA 算法的计算时间近似为 PSO 算法的一半，这是因为 QSHDGRA 算法可以利用商空间迅速收敛到最优值，因此计算时间可以保持近似线性增长。

### 4.5 与其他算法比较分析

将本文的 QSHDGRA 算法和 MinTime 最小时

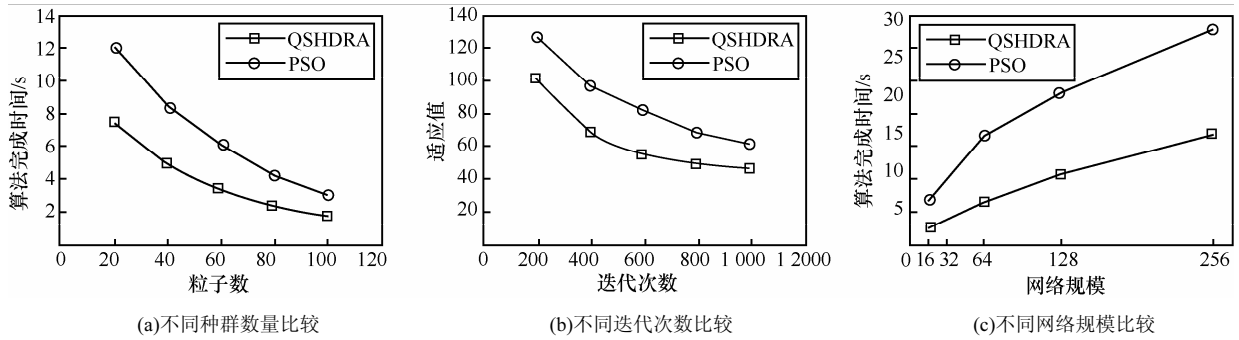


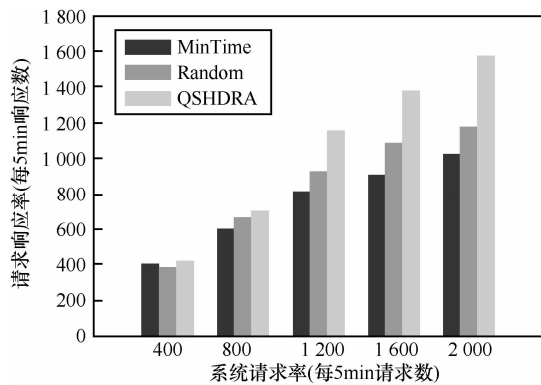
图 8 不同种群数量、迭代次数和网络规模下算法收敛速度比较

间算法、Random 算法的性能进行比较。MinTime 最小时间算法在搜索副本节点时仅仅以最快完成时间为目标，Random 算法则是随机从副本节点中选择一个节点获取数据。从图 9(a)和图 9(b)中可以看出，当请求率在 800 以下时，系统处于轻载，3 种算法的效率差别并不显著，QSHDGRA 算法的吞吐率略优于其他 2 种算法。随着请求率增大到 1 200 时，3 种算法效率出现明显差别，MinTime 算法最差，QSHDGRA 算法最优，而 Random 算法位于两者之间。此时由于系统尚未满载，三者的平均等待时间差别仍

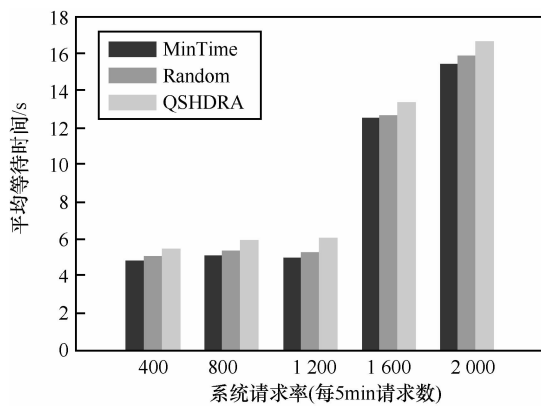
然近似相当。当请求率达到 2 000 时，系统处于重载时，QSHDGRA 和其他 2 种算法的差别进一步拉大。Random 算法之所以优于 MinTime 算法是因为随机选择进度可以在一定程度上对实现节点的负载均衡，而 QSHDGRA 算法由于兼顾了系统整体节点处理能力和网络能力的均衡性，虽然响应时间略大于其他 2 种算法，但吞吐率也显著优于其他 2 种算法。

### 5 结束语

本文针对传统数据网络调度算法在对层次式数据网络调度时出现的难以得到全局最优值和收敛速度过慢问题，提出了一种基于商空间的层次式数据网络资源调度算法。定义了基于不同 QoS 业务请求的平均等待时间和网络与节点资源利用均衡度的调和模型的目标函数，实现了基于商空间的层次式最优资源调度算法。仿真结果表明，该算法可以显著提升系统的吞吐率，加快收敛速度，并具备线性扩展能力。本文提出的算法虽然具备分布式调度的特性，但仍然属于一种全局控制算法。当全局调度节点不可用时，各节点如何分布协同实现资源调度将是下一步主要的研究方向。



(a)服务响应率比较



(b)平均等待时间比较

图 9 QSHDGRA 算法与其他算法性能比较

### 参考文献:

- [1] ZHANG J G, ZHANG K, YAN Y Y, *et al.* Grid-based implementation of XDS-I as part of image-enabled EHR for regional healthcare in Shanghai [J]. *International Journal of Computer Assisted Radiology and Surgery*, 2011, 6(2):273-284.
- [2] JBoss enterprise data Grid[EB/OL]. <http://www.jboss.com/edg6-early-access/>, 2012.
- [3] CHENG K Y, WANG H H, WEN C H, *et al.* Dynamic file replica location and selection strategy in data grids[A]. *Proceedings of the 1st IEEE International Conference on Ubi-Media Computing and Workshops[C]*. Lanzhou, China, 2008. 484-489.
- [4] AL-MISTARIHI H H E, YONG C H. On fairness, optimizing replica selection in data grids[J]. *IEEE Transactions on Parallel and*

- Distributed Systems, 2009, 20(8):1102-1111.
- [5] DU W, CUI G H, LIU W. Reliability-aware replica selection for data-intensive applications on data grids[J]. Information and Intelligent Interdisciplinary Journal, 2011, 14(12):3913-3920.
- [6] QU M C, WU X H, LIAO M H, *et al.* A novel resource selection model for data grid based on QoS[J]. New Trends and Applications of Computer-Aided Material and Engineering, 2011, 186:203-209.
- [7] TANG B, ZHANG L. Optimal replica selection algorithm in data grid[J]. Theoretical and Mathematical Foundations of Computer Science, 2011, 164:297-304.
- [8] RAHMAN R M, BARKER K, ALHAJJ R. A predictive technique for replica selection in grid environment[A]. Proceedings of the 7th IEEE International Symposium on Cluster Computing and the Grid[C]. Rio de Janeiro, Brazil, 2007. 163-170.
- [9] NASEERA S, MADHU MURTHY K V. Performance evaluation of predictive replica selection using neural network approaches[A]. Proceedings of 2009 International Conference on Intelligent Agent and Multi-Agent Systems[C]. Chennai, India, 2009.
- [10] RAHMAN R M, ALHAJJ R, BARKER K. Replica selection strategies in data grid[J]. Journal of Parallel and Distributed Computing, 2008, 68(12):1561-1574.
- [11] ALMUTTAIRI R M, WANKAR R, NEGI A, *et al.* Replica selection in data grids using preconditioning of decision attributes by K-means clustering (K-RSDG)[A]. Proceedings of the 2nd Vaagdevi International Conference on Information Technology for Real World Problems[C]. Warangal, India, 2010. 18-23.
- [12] LI J. A replica selection approach based on prediction in data grid[A]. Proceedings of the 3rd International Conference on Semantics, Knowledge, and Grid[C]. Xi'an, China, 2007. 274-277.
- [13] WU C, WU K G, CHEN M, *et al.* Dynamic replica selection services based on state evaluation strategy[A]. Proceedings of the 4th ChinaGrid Annual Conference[C]. Yantai, China, 2009. 116-119.
- [14] MUNOZ V M, VICENTE G A, CARBALLEIRA F G, *et al.* Emergent algorithms for replica location and selection in data grid[J]. Future Generation Computer Systems-the International Journal of Grid Computing-Theory Methods and Applications, 2010, 26(7):934-946.
- [15] XIONG R Q, LUO J Z, SONG A B, *et al.* QoS preference-aware replica selection strategy using mapreduce-based PGA in data grids[A]. Proceedings of the 40th International Conference on Parallel Processing[C]. 2011. 394-403.
- [16] WANG X, DING S F. An overview of quotient space theory[J]. Sports Materials, Modelling and Simulation, 2011, 187:326-331.

#### 作者简介:



夏纯中(1981-),男,江苏南京人,江苏大学博士生,主要研究方向为分布式系统资源调度。



宋顺林(1947-),男,江苏常州人,江苏大学教授、博士生导师,主要研究方向为信息系统集成。