

模糊选择消费代价函数动态资源分配算法研究*

万李, 杨杰

(湖南科技学院 计算机与通信工程系, 湖南 永州 425100)

摘要: 主要研究了资源分配基数,提出了一种新的基于选择消费代价网格资源分配算法。提出的方法首先采用模糊基数建立一个消费代价函数模型,给出了网络中带权值的多参数消费代价函数。实例验证表明,当用户需要从系统中获得资源时,采用该方法在网格规模非常大而且计算量较小的情况下,选择代价最小的方法给用户使用来分配资源,具有一定的实际应用价值。

关键词: 存储网格; 消费代价函数; 动态资源分配; 代价最小

中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2012)07-2683-03

doi:10.3969/j.issn.1001-3695.2012.07.077

Realization of fuzzy selection-based consumption cost function in dynamic resource allocation algorithm

WAN Li, YANG Jie

(Dept. Computer & Communication Engineering, Hunan University of Science & Engineering, Yongzhou Hunan 425100, China)

Abstract: This paper mainly studied the resource allocation base, and put forward a kind of new consumer choice cost based on grid resource allocation algorithm. This method firstly used fuzzy base to build a consumer cost function model, and gave the network weighted multi parameter consumption cost function, the example showed that, when the user needs to be obtained from the system resource, proposed method makes illegal in the grid size is very large and small amount of calculation under the condition of minimum cost, choice method for user to allocate resources, has a certain practical value.

Key words: storage grid; consumption cost function; dynamic resource allocation; cost minimum

0 引言

企业存储网格已经步入一种分布式计算环境,目的在于利用互联网把分散在不同地理位置的各种可用空闲资源整合起来,实现计算、存储、数据等资源的全面共享,最终实现网络虚拟环境中的资源共享和协同工作。在任何资源共享的环境中,一个基本的服务就是资源分配,当给出一个所需资源的描述,计算此次资源分配各种方案所花费的代价,最后选择一种代价最小的方案为用户分配资源。

在实际应用中,企业存储网格的资源具有分布性、动态性、异构性,在设计网格中间件时需要重点考虑。其中动态性是源于企业存储网格会出现大量的节点和边,从而导致其拓扑结构出现点覆盖、边交叉以及图形拥塞等现象,动态性不仅体现在系统中的节点和共享的资源,它们可以在任何时间加入或离开,更体现在资源状态的快速变化。对于网格中的动态资源,由于它们的属性值变化非常的频繁,使得传统在分布式网络中应用 DHT 或 Flooding 技术很难实现对于网格动态资源的分配。针对这一问题,本文提出一种基于市场经济消费代价的网格动态资源分配算法,提出多参数的消费代价函数,通过该算法能够有效地刻画企业存储网格中的局部聚集特性,从而实现网格动态资源的查找和分配。由于真实世界中的企业存储网格普遍具有一定的聚类特性^[4]和动态性,本算法简单、有效且

具有良好的适用性。

1 代价函数模型的建立

1.1 最优方案设计

企业存储网格特点之一是对存储资源的动态管理。文件数据作为网格资源和信息交换的主要载体进行分配和存储时,过程中应尽量避免切割,但在实际操作过程来看,切割在所难免,一般来讲,如式(1)所示的就是切割文件所带来的开销:

$$u_i = \begin{cases} 1 & \text{切割第 } i \text{ 个文件} \\ 0 & \text{不切割第 } i \text{ 个文件} \end{cases} \quad (1)$$

当 u 和 s 均为 0-1 变量,如式(2)(3)所示。

$$u = \begin{cases} 1 & \text{切割文件} \\ 0 & \text{不切割文件} \end{cases} \quad (2)$$

$$s = \begin{cases} 1 & \text{从可靠度高的节点中替换文件} \\ 0 & \text{不从可靠度高的节点中替换文件} \end{cases} \quad (3)$$

这些方案对应的消费代价函数值即为所求的最优方案。

1.2 消费代价函数的建立

图1所示的资源节点带权图提出一种消费代价函数,同时本文采用的是模糊综合评价方法,并用以下的参数来进行权衡:

a)带权路径长度。显然带权路径长度越小,这个指标是在资源分配时的首要参数,消费代价函数值越小。消费代价随

收稿日期: 2011-12-21; 修回日期: 2012-02-21 基金项目: 湖南省自然科学基金资助项目(11JJ6056)

作者简介: 万李(1983-),男,湖北洪湖人,讲师,硕士研究生,主要研究方向为图像处理和模式识别、智能计算(12428112@qq.com); 杨杰(1976-),男(瑶族),湖南永州人,讲师,硕士研究生,主要研究方向为计算机网络安全、智能计算。

着节点之间物理距离的增大而增大,随着两节点链路网络流量的增大而减小,取网络流量的倒数作为图中边上的权,从而可利用 Dijkstra 算法计算图中某一节点到其他所有节点的最短路径及路径长度,再按照路径长度采用某种排序算法(如快速排序)对这些节点进行排序。

b) 节点空余存储空间。这个指标是在资源分配时的基本条件。空余存储空间满足用户数时才能将数据进行存入。

c) 节点可靠度。节点可靠度指的是端端连接时给定的本节点与目标节点能正常连通的概率,反映了网络的连通性^[7],本文中节点的可靠度化简为无节点失效的等价问题,按照 Ke 和 Wang 提出的分解算法将节点可靠度进行定性描述并量化。根据上述参数描述,给出消费代价函数如式(4)所示:

$$\varphi = \begin{cases} w_1(x-x') + w_2(yy') + w_3z + c_1 \times u + c_2 \times sx \\ \min(T(g) = \frac{1}{n} \sum_{i=1}^n (x_i - x)), s \times t \times (x - x') \geq 0 \\ y - y' \geq 0 \end{cases} \quad (4)$$

其中: φ 表示消费代价函数; w_1, w_2, w_3 为权系数; x, y, z 分别表示节点空余存储空间大小、节点可靠度、带权路径长度; x', y' 分别表示用户所需存储空间、所需节点可靠度; $T(g)$ 为 G 网格的拥挤度。拥挤度越小,带权路径长度越短,网格节点可靠度、空余存储空间与用户需求越匹配,则消费代价函数值越小。

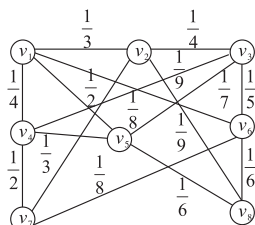


图1 资源节点带权示意图G

2 算法分析与设计

本文首先采用模糊基数建立一个消费代价函数模型,给出了网络中带权值的多参数消费代价函数。实例验证表明,当用户需要从系统中获得资源时,采用本文提出的方法在网格规模非常大而且计算量较小的情况下,得到最好的结果。

2.1 实例分析

首先制定用户的不同需求,如表1所示,分为用户所需数据的存储空间的大小,以及节点可靠度。

表1 用户需求

用户数据	指标	
	所需存储空间大小/GB	节点可靠度
D1	2.5	7
D2	3.0	5
D3	1.0	9

网格节点详细数据如表2所示。

表2 网格节点详细数据

网格节点	指标			
	空余存储大小/GB	可靠度	物理距离/m	平均网络流量/kbps
v_1	6.0	7	20	95
v_2	8.0	7	15	80
v_3	2.0	9	40	110
v_4	12.0	5	25	100
v_5	9.0	3	10	75
v_6	10.0	5	30	100
v_7	3.0	9	35	105
v_8	14.0	3	5	70

从表2可以看出,本文可以采用传统的 Dijkstra 算法求 G

中有效路径来保持畅通,即设某一顶点 v ,最短路径为 $P[v]$,带权路径长度为 $D[v]$,根据带权路径长度排序,依次可以得节点为 $v_{(1)}, v_{(2)}, \dots, v_{(n)}$,在存储空间和可靠度均满足 D1 的节点 v_1, v_2, v_7 ,有

$$\begin{matrix} \Delta x & \Delta y & z \\ v_1(N_1) & \begin{bmatrix} 3.5 & 0 & 0.0234 \\ 5.5 & 0 & 0.0208 \\ 0.5 & 2 & 0.0370 \end{bmatrix} \\ v_2(N_2) & \\ v_7(N_3) & \end{matrix}$$

记为 A (或 $(a_{ij})_{3 \times 3}$),其中的数字表示各个节点的取值。同时根据消费代价函数,对每一个节点进行判断。若在 $v_{(i)}$ 空余存储空间能满足用户需求,则数据存入节点。否则,如果 $v_{(i)}$ 可靠度高于用户需求,则将该节点做好记录为 $v_{(i)}$ 。

同时由于带权路径长度正比于物理距离,反比于平均网络流量,于是有

$$WPL = f_3 = z = \frac{\omega_1 a}{\omega_2 b} \quad (5)$$

选取 $\omega_1 = \frac{1}{10}, \omega_2 = \frac{9}{10}$, 即 $\frac{\omega_1}{\omega_2} = \frac{1}{9}$, 则表2可转换为表3所示。

表3 网格节点详细数据

网格节点	指标		
	空余存储空间大小/GB	可靠度	带权路径长度
v_1	6.0	7	0.023 4
v_2	8.0	7	0.020 8
v_3	2.0	9	0.040 4
v_4	12.0	5	0.027 8
v_5	9.0	3	0.014 8
v_6	10.0	5	0.033 3
v_7	3.0	9	0.037 0
v_8	14.0	3	0.007 9

由于每个影响因素的量纲不一样,故采用向量归一法进行标准化。

对于第一个指标 Δx , 令

$$\begin{cases} x_{11} = \frac{a_{11}}{\sqrt{\sum_{i=1}^3 a_{i1}^2}} = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2 + a_{31}^2}} \\ x_{21} = \frac{a_{21}}{\sqrt{\sum_{i=1}^3 a_{i1}^2}} = \frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2 + a_{31}^2}} \\ x_{31} = \frac{a_{31}}{\sqrt{\sum_{i=1}^3 a_{i1}^2}} = \frac{a_{31}}{\sqrt{a_{11}^2 + a_{21}^2 + a_{31}^2}} \end{cases} \quad (6)$$

式(6)中 x_{11} 表示经过标准化后,节点 v_1 在指标 Δx 上的值,即节点 v_1 空余存储空间与用户所需存储空间的差; x_{21} 表示经过标准化后,节点 v_2 在指标 Δx 上的值,即 v_2 空余存储空间与用户所需存储空间的差; x_{31} 表示经过标准化后,节点 v_7 (亦即 N_3) 在指标 Δx 上的值,即 v_7 空余存储空间与用户所需存储空间的差。

对于第二个指标 Δy , 令

$$\begin{cases} x_{12} = \frac{a_{12}}{\sqrt{\sum_{i=1}^3 a_{i2}^2}} = \frac{a_{12}}{\sqrt{a_{12}^2 + a_{22}^2 + a_{32}^2}} \\ x_{22} = \frac{a_{22}}{\sqrt{\sum_{i=1}^3 a_{i2}^2}} = \frac{a_{22}}{\sqrt{a_{12}^2 + a_{22}^2 + a_{32}^2}} \\ x_{32} = \frac{a_{32}}{\sqrt{\sum_{i=1}^3 a_{i2}^2}} = \frac{a_{32}}{\sqrt{a_{12}^2 + a_{22}^2 + a_{32}^2}} \end{cases} \quad (7)$$

式(7)中 x_{12} 表示经过标准化后,节点 v_1 在指标 Δy 上的

值,即节点 v_1 可靠度与用户要求可靠度的差; x_{22} 表示经过标准化后,节点 v_2 在指标 Δy 上的值,即 v_2 可靠度与用户要求可靠度的差;同理, x_{32} 也表示类似含义。

对于第三个指标 z (或 f_3),令

$$\begin{cases} x_{13} = \frac{a_{13}}{\sqrt{\sum_{i=1}^3 a_{i3}^2}} = \frac{a_{13}}{\sqrt{a_{13}^2 + a_{23}^2 + a_{33}^2}} \\ x_{23} = \frac{a_{23}}{\sqrt{\sum_{i=1}^3 a_{i3}^2}} = \frac{a_{23}}{\sqrt{a_{13}^2 + a_{23}^2 + a_{33}^2}} \\ x_{33} = \frac{a_{33}}{\sqrt{\sum_{i=1}^3 a_{i3}^2}} = \frac{a_{33}}{\sqrt{a_{13}^2 + a_{23}^2 + a_{33}^2}} \end{cases} \quad (8)$$

式(8)中 x_{13} 表示经过标准化后,节点 v_1 在指标 z 上的值,即 v_1 的带权路径长度; x_{23} 表示经过标准化后,节点 v_2 在指标 z 上的值,即 v_2 的带权路径长度;同理, x_{33} 也表示类似含义。

联立式(6)~(8)可得

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 0.535 & 0 & 0.482 \\ 0.841 & 0 & 0.429 \\ 0.076 & 1 & 0.763 \end{bmatrix} \quad (9)$$

由于每一个影响因素对消费代价函数 φ 的影响程度有些差异,因而进一步采用如式(10)的相对比较法确定权重:

$$h_{ij} = \begin{cases} 1 & \text{当 } f_i \text{ 比 } f_j \text{ 重要时} \\ 0.5 & \text{当 } f_i \text{ 与 } f_j \text{ 同等重要时 } (i, j = 1, 2, 3) \\ 0 & \text{当 } f_i \text{ 不如 } f_j \text{ 重要时} \end{cases} \quad (10)$$

其中: f_1 对应本例中 Δx ; f_2 对应本例中 Δy ; f_3 即为本例中 z :

$$H = (h_{ij})_{3 \times 3} \quad i, j = 1, 2, 3 \quad (11)$$

显然,

$$\begin{cases} h_{ii} = 0.5 \\ h_{ij} + h_{ji} = 1 \end{cases} \quad i, j = 1, 2, 3 \quad (12)$$

根据具体设计的要求,可以凭经验事先定性为影响消费代价函数的各因素(即评价指标 f_1, f_2, f_3)赋权值,如 Δy 的权值 $> z$ 的权值 $> \Delta x$ 的权值,则可得到

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 \\ 1 & 0.5 & 1 \\ 1 & 0 & 0.5 \end{bmatrix} \quad (13)$$

又有

$$w_i = \frac{\sum_{j=1}^3 h_{ij}}{\sum_{i=1}^3 (\sum_{j=1}^3 h_{ij})} \quad (i = 1, 2, 3) \quad (14)$$

联立(13)(14)式解得

$$w_1 = \frac{h_{11} + h_{12} + h_{13}}{h_{11} + h_{12} + h_{13} + h_{21} + h_{22} + h_{23} + h_{31} + h_{32} + h_{33}} \cdot \frac{1}{9} \quad (15)$$

同理,

$$w_2 = \frac{h_{21} + h_{22} + h_{23}}{h_{11} + h_{12} + h_{13} + h_{21} + h_{22} + h_{23} + h_{31} + h_{32} + h_{33}} = \frac{5}{9} \quad (16)$$

$$w_3 = \frac{h_{31} + h_{32} + h_{33}}{h_{11} + h_{12} + h_{13} + h_{21} + h_{22} + h_{23} + h_{31} + h_{32} + h_{33}} = \frac{1}{3} \quad (17)$$

从而将定性描述的权值转换成了定量的权值。

由数据 D_1 的各项要求可知,既不需要切割文件,也不需要从网格中可靠度高的节点中替换老的、不太重要的数据,再由式(2)(3)知,0-1 变量 u, s 均为零,从而 $c_1 \times u, c_2 \times s$ 对消费代价函数 φ 的贡献均为零, φ 值只取决于前三项。根据线性加

权和法,节点 v_1, v_2, v_7 消费代价函数值为

$$\begin{cases} \varphi_1 = w_1 x_{11} + w_2 x_{12} + w_3 x_{13} \\ \varphi_2 = w_1 x_{21} + w_2 x_{22} + w_3 x_{23} \\ \varphi_7 = w_1 x_{31} + w_2 x_{32} + w_3 x_{33} \end{cases} \quad (18)$$

联立式(14)~(18)可得:

$$\begin{cases} \varphi_1 \approx 0.220 \\ \varphi_2 \approx 0.236 \\ \varphi_7 \approx 0.818 \end{cases} \quad (19)$$

经计算 $\varphi_1 < \varphi_2 < \varphi_7$,故节点 v_1 的消费代价函数值最小,节点 v_2 次之,节点 v_7 的最高,因此, v_1 为最佳目标节点,即最优解。

2.2 实验结果与结论

消费代价函数最小者,也就是说第 t 个被记录的节点为所求目标节点,同时还应用到了替换算法。为了验证本算法的有效性,本文在 Windows XP 操作平台上,使用 Eclipse 插件开发技术开发了一个自动生成企业存储网络图的软件,能自动生成全局和局部拓扑图。为了使测试数据达到一定规模,选用了 Karate 和 Dolphins 网络进行研究。实验表明,本算法适用于处理规模在 4 000 个节点以下的企业存储网络,尤其在在 600 个以下节点的时候,能稳定在 5 s 左右计算出最小的消费代价函数,给出最佳资源分配方案。

3 结束语

本文提出了一种基于市场经济消费代价的网络动态资源分配算法。该算法利用模糊选择方法来建立一个消费代价函数计算模型,针对动态网络的拓扑图中出现点覆盖、边交叉以及图形拥塞、网络中的节点往往具有不确定性等现象给出带权值的多参数消费代价函数,和实际数学模型及应用实例。仿真实验结果表明,在企业存储网络模型动态化情形下,当用户需要在系统中获取资源,通过本算法能在网络规模较大且计算量较少的情况下选择一种代价最小的方案给用户分配资源,具实际的应用价值。

参考文献:

- [1] WALTER F E, BATI' IS, IION S, SCHWERTZER F. A model of a trust-based recommendation system on a social network[J]. *Autonomous Agents and Multi-agents Systems*, 2006, 16(1): 57-74.
- [2] WU Hui-ju, TING I H, WANG Kai-yu. Combining social network analysis and Web mining techniques to discover interest groups in the blogspace[C]//Proc of the 4th International Conference on Innovative Computing Intemation and Control. 2009:1180-1183.
- [3] 李卓远,吴为民,洪先龙. 优化线长和拥挤度的增量式布局算法[J]. *计算机辅助设计与图形学学报*, 2003, 15(6): 651-655.
- [4] Von LANDEBERGER T, KUIJPER A, SCHRECK T, et al. Visual analysis of Large graphs[J]. *IEEE Trans on Visualization and Computer Graphics*, 2010, 29(2): 37-60.
- [5] 曾春,邢春晓,周立柱. 个性化服务技术综述[J]. *软件学报*, 2002, 13(10): 1952-1961.
- [6] HERLOCKER J, KONSTAN J. Evaluating collaborative filtering recommender systems[J]. *ACM Trans on Information Systems*, 2004, 22(1): 5-53.
- [7] 张波,雋志才,林徐勳. 基于有限理性的弹性需求随机用户均衡交通分配模型[J]. *计算机应用研究*, 2011, 28(9): 3268-3271.