

基于 Biswapped 网络混合有效的负载均衡算法研究

童朝南, 孙丽婷

(1. 北京科技大学自动化学院, 北京 100083; 2. 北京科技大学钢铁流程先进控制教育部重点实验室, 北京 100083)

摘要: 本文针对大型层次网络 Biswapped Network(BSN), 提出了一种简单有效的负载均衡策略 CDE-X, 克服了传统方案中因大规模层次网络复杂高阶的拉普拉斯矩阵特征值的计算而呈现的不适用性缺点. 在新算法中, 均衡迭代过程大部分发生在 BSN 的因子网络内部, 因而, CDE-X 只需知道因子网络的结构及拉普拉斯特征值, 从而在迭代均衡过程中提高了迭代收敛速度, 并且降低了迁移负载流量的计算复杂性, 同时降低了通信流量. 理论比较证明, CDE-X 不止降低了计算复杂性, 更加降低了均衡迭代的步数, 比传统的负载均衡策略 X 要更加简单、有效, 更适用于大规模层次网络 BSN.

关键词: Biswapped Network (BSN); 拉普拉斯矩阵; 特征值; 迭代

中图分类号: TP301 **文献标识码:** A **文章编号:** 0372-2112 (2013) 08-1529-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2013.08.012

A Hybrid Efficient Load Balancing Algorithm on Biswapped Network

TONG Chao-nan, SUN Li-ting

(1. School of Automation, University of Science and Technology Beijing, Beijing 100083, China;

2. Key Laboratory of Advanced Control of Iron and Steel Process (Ministry of Education), Beijing 100083, China)

Abstract: In view of the large hierarchical Network Biswapped Network(BSN), paper suggests a simple and effective load balancing strategy CDE-X. This new scheme overcomes the unworthiness of the large-scale hierarchical network's traditional scheme in which the calculation of the complex high order Laplacian matrix's eigenvalues. In the new method, most load balancing iterations are carried out in the factor network, therefore, CDE-X only needs to know the factor network's structure and the Laplacian eigenvalues, which improves the iterative convergence speed in the process of iterative balancing, and reduces the calculation complexity of the flows, and reduce the communication flows. According to the comparison theory, CDE-X not only reduces the computational complexity, but also reduces the iteration steps, which is more simple and effective than the traditional strategy X, and more applicable to the large scale hierarchical network BSN.

Key words: Biswapped Network(BSN); Laplacian matrix; eigenvalues; iteration

1 引言

在并行计算中, 动态的不可预测的任务特性往往会引起系统内不均衡的负载分布, 从而造成整个系统的性能下降, 因而, 如何使并行计算机系统的利用率达到最大, 负载均衡变的尤为关键^[1,2].

目前, 已经有许多建立在普遍网络基础上的负载均衡算法被提出. 针对同构系统, Cybenko 等^[3]提出了局部扩散算法的一般形式. Muthukrishnan^[4]在 Cybenko 的基础上提出了 Second Order Scheme(SOS), 加快了迭代的速度. Diekmann 等^[5]设计了一种新的迭代算法, 称之为 the

Optimal Polynomial Scheme(OPS), 该算法更加提升了迭代过程的收敛速度, 降低了迭代步数. Elsasser 等^[6]介绍了 optimal diffusion scheme(OPT), 虽然迭代步数与 OPS 同样, 但与 OPS 对比, OPT 方法更加的简单、有效. 而针对异构系统, Elsass 等^[7]提出了扩展的均衡模型. Rotaru 等^[8]提出了从同构系统到异构环境的过渡. 他们都在理论上证明了算法的有效性. 然而, 因为拉普拉斯矩阵特征值复杂的计算性, 以上传统的扩散算法并不适用于大规模的层次网络. 因此, 基于大型层次网络的负载均衡算法的研究是非常必需的^[9].

目前存在的大规模的层次网络系统, 如 OTIS 网络

(Optical Transpose Interconnection System)^[10,11] 及 Biswapped 网络(BSN),都是紧密相关的.但因 BSN 的低扩展成本、对称性及可以继承因子网络的众多理想特性等优势,吸引了越来越多的学者研究^[12,13].因而,如何利用 BSN 的高扩展性和容错性来提高并行计算机系统的负载均衡问题成为了研究的热点.

2 相关概念和性质

定义 1 (Biswapped 网络(BSN)).假设 $\Omega = (V_\Omega, E_\Omega)$ 是一个无向连通图, $|V(\Omega)| = n$.以 Ω 为因子网络构建的 BSN 同样是一个无向图,记作 $\text{BSN}(\Omega)$,其顶点集可表示为

$$V(\text{BSN}(\Omega)) = \{ \langle i, g, p \rangle \mid g, p \in V(\Omega), i = 0, 1 \},$$

$$\text{边集表示为}$$

$$E(\text{BSN}(\Omega)) = \{ (\langle i, g, p_1 \rangle, \langle i, g, p_2 \rangle) \mid (p_1, p_2) \in E(\Omega), i = 0, 1 \} \cup \{ (\langle i, g, p \rangle, \langle 1-i, p, g \rangle) \mid g, p \in V(\Omega), i = 0, 1 \}.$$

由定义可以得出, $\text{BSN}(\Omega)$ 被平均分成两部分:部 0 和部 1.另外,它是由 $2n$ 个点不相交子网络 Ω (被称作组或簇)组成.如果 Ω 有 n 个节点,则 $\text{BSN}(\Omega)$ 则拥有 $2n^2$ 个节点.

在 $\text{BSN}(\Omega)$ 中,节点 $\langle i, g, p \rangle$ 表示第 i 部中的组 g 中的处理器 p .在网络内部,组间链将一部内的组 g 中的处理器 p 与另一部内的组 p 中的处理器 g 相连^[12].因 BSN 分为两部,且两部间有链接,因而,如果把 BSN 中的每个因子网络看作一个超点,则 $\text{BSN}(\Omega)$ 可看作一个完全两部图 K_n, n .

针对第 i 部的第 j 组的因子网络, Ω_{ij} 的顶点集合也可表示为 $V_{ij} = \{ v_{ik} \mid 0 \leq k \leq n-1, i = 0, 1 \}$, 边的集合 $E_{ij} = \{ (v_{im}, v_{jn}) \mid (v_m, v_n) \in E_\Omega, i = 0, 1 \}$.由此, $\text{BSN}(\Omega)$ 的顶点集合 $V_\Omega = \bigcup_{0 \leq j \leq n-1, i=0,1} V_{ij}$, 边的集合可以分成两个子集 E_1 和 E_2 : E_1 代表了组内的链接,表示成 $E_1 = \bigcup_{i=0,1; 0 \leq j \leq n-1} E_{ij}$; E_2 表示部间链接,表示成 $E_2 = \{ \langle v_{ij}, v_{(1-i)kj} \rangle \mid k < j, i = 0, 1 \}$.

图 1 给出了以环网络 R_4 为因子网络的 BSN.

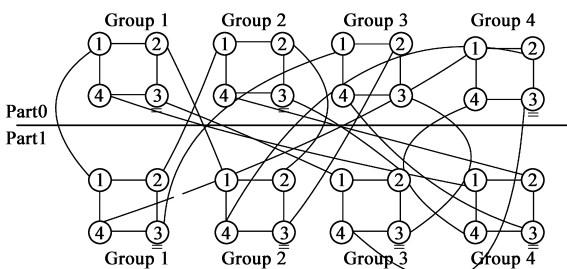


图 1 BSN- R_4 (R_4 是四个节点的环)

假设 $W_{ij} = (w_{ij1}, w_{ij2}, \dots, w_{ijn})^T (i = 0, 1)$, $C_{ij} = (c_{ij1},$

$c_{ij2}, \dots, c_{ijn})^T (i = 0, 1)$ 表示 Ω_{ij} 的负载和权值; $W_i = (w_{i1}, w_{i2}, \dots, w_{in})^T (i = 0, 1)$, $C_i = (c_{i1}, c_{i2}, \dots, c_{in})^T (i = 0, 1)$ 表示 $\text{BSN}(\Omega)$ 的负载和权值.其中,定义 Ω_{ij} 的第 k 个节点 v_{ijk} 的初始负载 $w_{ijk}^0 \geq 0$, 初始权值 $c_{ijk} > 0$.另外,定义对角矩阵 C' 和 C'_{ij} , 其对角线元素为向量 C_i 和向量 C_{ij} , 可表示为

$$C' = \text{diag}(c_{011}, c_{012}, \dots, c_{01n}, c_{021}, \dots, c_{02n}, \dots, c_{0m1}, c_{111}, \dots, c_{1m1})$$

$$C'_{ij} = \text{diag}(c_{0i1}, c_{0i2}, \dots, c_{0in}, c_{1i1}, c_{1i2}, \dots, c_{1in}) (i = 0, 1).$$

假设 A 及 A^f 分别为 $\text{BSN}(\Omega)$ 和因子网络 Ω 的定向关联矩阵,而 A_2 定义为是 E_2 与 $\text{BSN}(\Omega)$ 相关节点关联的矩阵.其中,矩阵 A, A^f 及 A_2 的每一列恰有两个非零元,一个是 1,另一个为 -1 ,分别代表节点相关边的定向.

图的拉普拉斯矩阵可表示为 $L = AA^T$. 本文分别用 L 及 L^f 表示 $\text{BSN}(\Omega)$ 及因子网络 Ω 的拉普拉斯矩阵.假设 L 及 L^f 的特征值分别用 λ_i 及 λ_i^f 表示,特征值依次递减排序为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 及 $\lambda_1^f \geq \lambda_2^f \geq \dots \geq \lambda_m^f$, 并且对应的特征向量分别用 z_i 及 z_i^f 来表示.假使对于 $\text{BSN}(\Omega)$ 网络上的所有边的边权值 α_e 都赋予一个固定的值 $\alpha \in (0, 1)$, Ω 网络上边的边权值 α_e^f 都赋予一个固定的值 $\alpha^f \in (0, 1)$, 则在同构网络系统中,边权值向量内的元素都是相同的,并且由此所对应的扩散矩阵 M 以及 M^f 可表示为 $M = I_{2n} - \alpha L$ 以及 $M^f = I_n - \alpha^f L^f$.那么 M 以及 M^f 所分别对应的特征值为 $\mu_i = 1 - \alpha \lambda_i$, $\mu_i^f = 1 - \alpha^f \lambda_i^f$.另外,用 γ 以及 γ^f 表示矩阵 M 以及 M^f 的第二大特征值的绝对值,即 $\gamma = \{ \max(|\mu_2|, |\mu_m|) \} < 1$, $\gamma^f = \{ \max(|\mu_2^f|, |\mu_m^f|) \} < 1$.它们被称为是扩散矩阵的扩散范式. γ 值越小,网络系统的收敛速度就会越快.

在基于多项式的扩散模型方案中,工作负载 w^k 在第 k 步迭代之后的表达式可表示为 $w^k = p_k(M)w^0$. 多项式 p_k 的收敛依赖于误差 $e^k = w^k - \bar{w}$ 是否趋于 0. 误差项 e^k 是衡量经过 k 步迭代后负载均衡效果的重要指标,根据文献[3],误差项 e^k 满足 $e^k = p_k(M)e^0$, 同时知道 $\|e^k\|_2 \leq \max_{i=2}^m |p_k(\mu_i)| \cdot \|e^0\|_2, k = 0, 1, \dots$.

在异构的网络系统中, $\text{BSN}(\Omega)$ 及因子网络 Ω 的关联矩阵 \tilde{A} 以及 \tilde{A}^f 可分别表示为 $\tilde{A} = AF^{-1}$, 及 $\tilde{A}^f = A(F^f)^{-1}$. 其中矩阵 F 为对角矩阵, 对角元素 $F_{ii} = \sqrt{f_{ii}}$ (f_{ii} 表示 $\text{BSN}(\Omega)$ 边 i 的边权值); 矩阵 F^f 的对角元素 $F_{ii}^f = \sqrt{f_{ii}^f}$ (f_{ii}^f 表示 Ω 中边 i 的边权值). 那么 $\text{BSN}(\Omega)$ 及因子网络 Ω 的拉普拉斯矩阵表达式的一般形式分别为 $\tilde{L}C^{-1}$ 及 $\tilde{L}^f C_{ij}^{-1}$ 且 $\tilde{L} = \tilde{A}\tilde{A}^T, \tilde{L}^f = \tilde{A}^f \tilde{A}^{fT}$. 其实,异构网络中的特征值分析与同构网络架构下的分析是相同

的,而不同的一点即是研究对象由 $\tilde{L}C^{-1}$ 替代了 L .

3 BSN 中的负载均衡

由上述内容可知,在负载均衡策略研究中,拉普拉斯及扩散矩阵研究是非常关键的部分,但要计算大规模层次网络的拉普拉斯及扩散矩阵的特征值及特征向量是非常消耗计算时间并且是非常复杂的.因此,针对大规模的 BSN,为了降低此类复杂矩阵的特征值及特征向量的计算复杂性,本文提出一种新型的混合、有效的可适用性负载均衡方案.

3.1 同构的 BSN 下的负载均衡策略

针对同构的 BSN(Ω),本文提出一种混合的负载均衡方案 CDE-X (Combination of Diffusion and Exchange scheme)(X 代表传统的策略方案,例如 SOS, FOS, OPT 或 OPS 等),方案的基本思想就是把整个负载均衡过程划分成四个阶段,即 Diffusion, Exchange, Diffusion, Exchange. 其策略步骤如下:

Step1 (Diffusion)^[14]: 经过局部迭代的负载均衡过程,BSN(Ω)的每个因子网络 Ω_{ij} 内部节点上的负载都变的相同,因子网络内部达到局部均衡状态.

在第一阶段过程中,经过 k 步迭代过程,每个因子网络中的负载向量为 $W_{ij}^k = (In \otimes p_k(M^f)) W_{ij}^0$.

假设 y_f 表示在第 k 步迭代过程中通过因子网络边 e 迁移的负载量,而 x_f 表示经过 k 步迭代过程之后通过边 e 迁移的所有负载流量.其中,流量方向取决于对应关联矩阵中边的指向;并且,流量 x 若满足 $Ax = w^0 - \bar{w}$,则 x 称之为均衡流量.

经过 k 步迭代,可以计算出经过第一阶段(Diffusion),每个因子网络可产生的流量 x_Ω 为^[15]:

$$\begin{aligned} x_\Omega &= x_f = x^{k-1} + \alpha^f (A^f)^T W_{ij}^{k-1} \\ &= \alpha^f \sum_{p=0}^{k-1} (A^f)^T (M^f)^p W_{ij}^0 \\ &= \alpha^f \sum_{p=0}^{k-1} (A^f)^T (M^f)^p \sum_{i=2}^m z_i^f \\ &= \alpha^f \sum_{p=0}^{k-1} (A^f)^T \sum_{i=2}^m (\mu_i^f)^p z_i^f \\ &= \alpha^f (A^f)^T \sum_{i=2}^m \sum_{p=0}^{k-1} (\mu_i^f)^p z_i^f \\ &= \alpha^f (A^f)^T \sum_{i=2}^m \frac{1 - (\mu_i^f)^k}{1 - \mu_i^f} z_i^f \end{aligned}$$

由 $\mu_i^f = 1 - \alpha^f \lambda_i^f$, 上式可转换为

$$x_\Omega = x_f^k = (A^f)^T \sum_{i=2}^m \frac{1 - (\mu_i^f)^k}{\lambda_i^f} z_i^f = (A^f)^T \sum_{i=2}^m \frac{1}{\lambda_i^f} z_i^f$$

推算到整个 BSN,则经过第一阶段产生的流量迁移总量 x_1 则可表示为:

$$x_1 = \sum_{j=1}^{2n} x_{\Omega,j} = (I \otimes (A^f)^T) \sum_{j=1}^{2n} \sum_{i=2}^m \frac{1}{\lambda_{i,j}^f} z_{i,j}^f \quad (1)$$

Step2 (Exchange): 利用维交换算法来互换所有节点对 $\langle i, g, p \rangle$ 与 $\langle 1-i, p, g \rangle$ 上的负载.

经过此过程,通过部间链接产生的迁移流量则可以表示为^[15]

$$x_2 = (I \otimes A_2^T) \sum_{i=2}^m \frac{1}{\lambda_i^f} z_i^f \quad (2)$$

经过此次交换负载后,每个因子网络的负载总值达成一致.

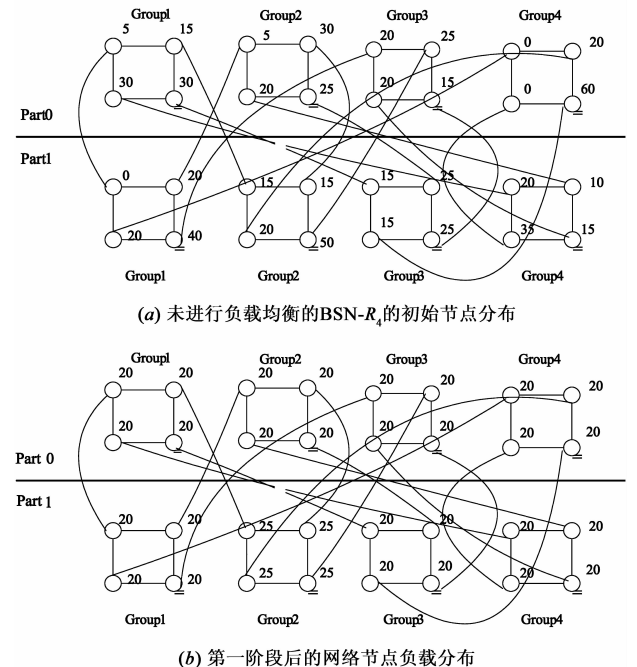
Step3 (Diffusion): 依然运用扩散算法,在经过低于 m^f 次的迭代过程之后,每个因子网络内部再次达到局部负载均衡状态.计算均衡流量的方法与 **Step1** 相同,且 x_3 可表示为:

$$x_3 = (I \otimes (A^f)^T) \sum_{j=1}^{2n} \sum_{i=2}^m \frac{1}{\lambda_{i,j}^f} z_{i,j}^{f*} \quad (3)$$

Step4 (Exchange): 与 **Step2** 不同的是,对每对 $\langle i, g, p \rangle$ 与 $\langle 1-i, p, g \rangle$ 逐一进行交换扩散.经过第 k 步迭代,边 e 上的负载迁移量为 $y_e^k = 1/2k (w_{ij}^k - w_{(1-i)j}^k)$. 经过此阶段迭代过程, w_{ijk} 将收敛于整个网络节点负载的均值,BSN(Ω)也会达到全局的均衡状态.在此阶段中,产生的流量 x_4 可表示为:

$$x_4 = (I \otimes A_2^T) \sum_{i=2}^m \frac{1}{\lambda_i^f} z_i^{f*} \quad (4)$$

图 2 给出了以环网络 R_4 为因子网络的同构 BSN- R_4 的负载均衡过程范例.



(b) 第一阶段后的网络节点负载分布

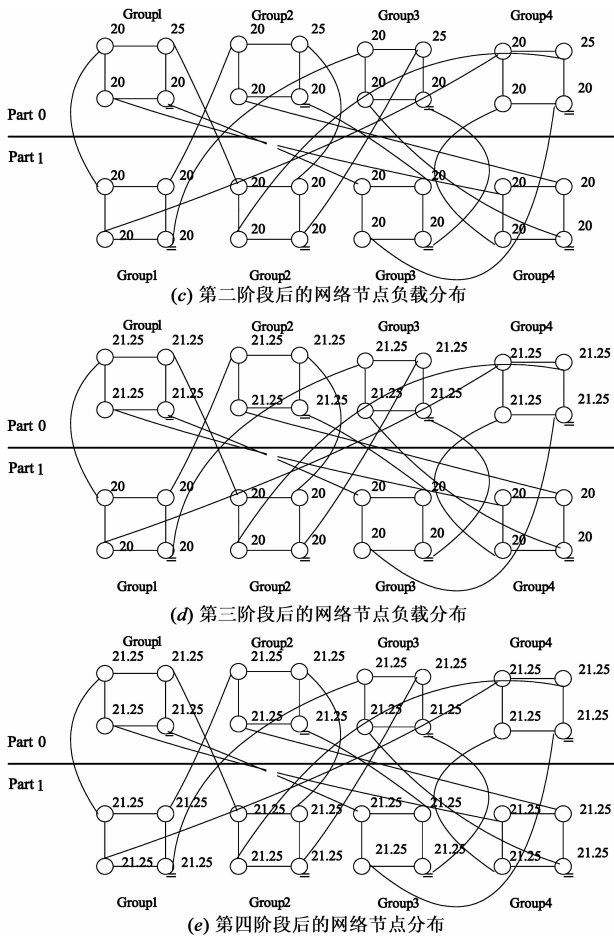


图2 CDE-X应用于同构的BSN-R_n上的负载均衡示例 (节点旁的数字代表节点负载)

通过以上策略分析可得,如果把传统的负载均衡模型 X 作用于 BSN(Ω),则经过 k 步负载均衡过程后的迁移总量 x 可表示为

$$\begin{aligned}
 x &= (I \otimes A^T) \sum_{i=2}^m \frac{1}{\lambda_i} z_i \\
 &= I \otimes [A_2 A^{f_1} A^{f_2} \cdots A^{f_{2n}}]^T \sum_{i=2}^m \frac{1}{\lambda_i} z_i \quad (5)
 \end{aligned}$$

而当运用 CDE-X 算法时,整个网络产生的迁移流量 x' 则可表示为

$$\begin{aligned}
 x' &= x_1 + x_2 + x_3 + x_4 \\
 &= (I \otimes (A^f)^T) \sum_{j=1}^{2n} \sum_{i=2}^{m'} \frac{1}{\lambda_{i,j}^f} z_{i,j}^f \\
 &+ (I \otimes A_2^T) \sum_{i=2}^{m_2} \frac{1}{\lambda_i^f} z_i^f \\
 &+ (I \otimes (A^f)^T) \sum_{j=1}^{2n} \sum_{i=2}^{m'} \frac{1}{\lambda_{i,j}^f} z_{i,j}^{f*} \\
 &+ (I \otimes A_2^T) \sum_{i=2}^{m_2} \frac{1}{\lambda_i^*} z_i^* \\
 &= (I \otimes (A^f)^T) \sum_{j=1}^{2n} \sum_{i=2}^{m'} \frac{1}{\lambda_{i,j}^f} (z_{i,j}^f + z_{i,j}^{f*})
 \end{aligned}$$

$$+ (I \otimes A_2^T) \sum_{i=2}^{m_2} \frac{1}{\lambda_i^*} (z_i^* + z_i^f) \quad (6)$$

由式(5)可以了解,当作用一般模型 X 于 BSN 时,所产生的负载迁移总量计算必须提前得知整个网络系统的全局谱信息,并且必须计算出 $2n^2 \times 2n^2$ 矩阵的拉普拉斯特征值及特征向量,经研究表明,计算高阶矩阵的特征值一般比较困难.若因子网络又是由直径稍大的网络组成,则计算全局的谱值则是非常复杂而又不易的.然而,分析式(6)可知,CDE-X 方法的主要计算量都集中在因子网络内部,即,只需了解 Ω 的内部谱信息,也就是对应 $n \times n$ 阶拉普拉斯矩阵的特征值,并且,矩阵 A₂ 的计算量只达到数量级 O(n).

因此,根据式(5)及式(6)可看出,对比流量 x 的计算,运用 X 模型时,计算量可以达到 O(n³),而 CDE-X 策略的计算量最高达到 O(n²),从而降低了系统的计算复杂度;并且,对比 X 模型可知,在 CDE-X 方法中,BSN 中的负载迁移大部分都发生在因子网络内部,而因为因子网络内部的链接都是短距离的通信链,而因子网络间的链接都是较长距离的通信链^[12],因而,可知 CDE-X 模型方法大大减少了通信代价,降低了通信成本.

3.2 异构的 BSN(Ω)下的负载均衡策略

针对异构形态的 BSN(Ω),有以下需要注意的地方:

首先,由异构网络 BSN(Ω)的拉普拉斯特征值计算公式可知,网络中的每条边都被赋予一个边权值.

在本算法中,因子网络中每条边的权值 c_{ij} 可表示为 $c_{ij} = \frac{c_i \cdot c_j}{\sum_{k=1}^n c_k}$,其中, v_i 及 v_j 表示与边 e_{ij} 相关的两个节点.

对于任意组间链接 (< i, g, p >, < 1 - i, p, g >) 的边权值 l_{gp} 的表达式则有:

$$l_{gp} = \frac{(\sum_{i=1}^n c_{gi})(\sum_{i=1}^n c_{pi})}{(\sum_{i=1}^n \sum_{j=1}^n c_{ij})_0 + (\sum_{i=1}^n \sum_{j=1}^n c_{ij})_1}$$

在对异构系统实施负载均衡方案时,其对应的扩散和交换方案也都要进行相应的修正.

在 Step1 中,经过 k 步迭代过程,每个因子网络中的负载向量为 $W_{ij}^k = (I_n \otimes p_k (I - \alpha A (F^f)^{-1} ((F^f)^{-1})^T C_{ij}^{-1})) W_{ij}^0$;而在每个因子网络内部,当迭代结束后,每个节点的负载权值可表示为 $w_{ijk} = \frac{c_k}{\sum_{m=1}^n c_m} \sum_{l=1}^n w_l$;其迁移流量 $x_f^k = (A (F^f)^{-1})^T \sum_{i=2}^{m'} \frac{1}{\lambda_i^f} z_i^f$.

在 Step2 中,通过维交换过程,每条链接 l_{gp} 上迁移的流量为 $x_{gp} = \left| \frac{1_{gp}}{c_{gp}} \times w_{gp} - \frac{1_{gp}}{c_{pg}} \times w_{pg} \right|$, 流量迁移方向由差值的正负决定.

Step3 与 Step4 则将与上述 Step1 及 Step2 相类似.

由此,通过上述四个阶段的迭代过程,整个异构系统的单位节点的负载都将与其自身性能达成优化的比例,负载均衡效果显著.

针对异构的 BSN,对比 X 模型及 CDE-X 可知,CDE-X 更具有适用性,其分析过程与 3.1 节相类似,因而,此处略去说明.

图 3 给出了异构环境下 BSN- R_4 的负载均衡过程范例,算法实现参见下文.

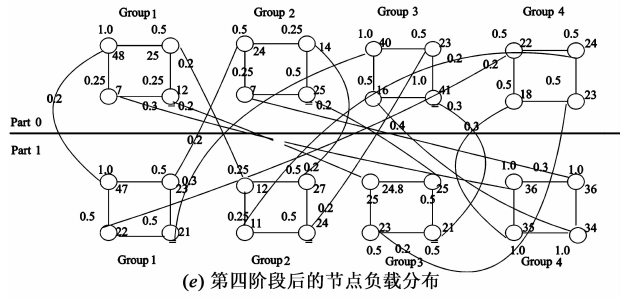
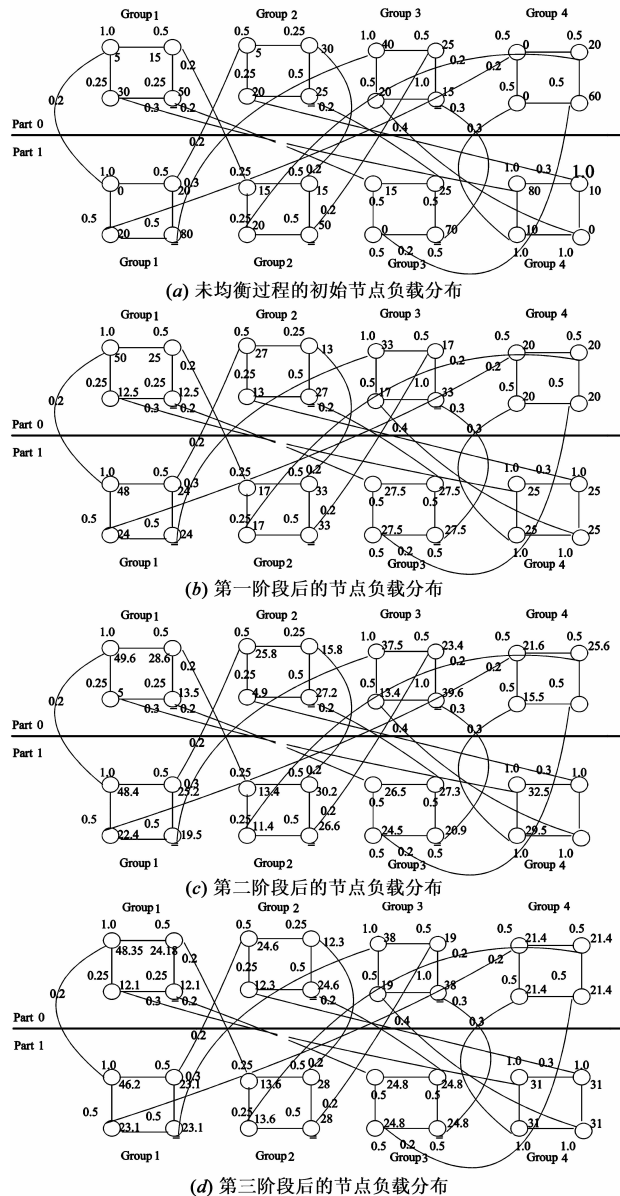


图3 CDE-X应用于异构的BSN- R_4 上的负载均衡示例(节点旁的较小数字代表节点权值,链路上的数字代表边权值)

CDE-X 算法

Algorithm CDE-X for load balancing

Require: Factor network $\Omega_{ij}, W_{ij}, C_{ij}$;

Ensure: Balancing load vector w_{ij}^l ;

- 1 for all groups Ω_{ij} of BSN(Ω);
- 2 run the Diffusion procedure on Ω_{ij} ;
- 3 end for;
- 4 for all intergroup edges $e: (\langle i, g, p \rangle, \langle 1-i, p, g \rangle)$;
- 5
$$l_{gp} = \frac{(\sum_{i=1}^n c_{gi})(\sum_{i=1}^n c_{pi})}{(\sum_{i=1}^n \sum_{j=1}^n c_{ij})0 + (\sum_{i=1}^n \sum_{j=1}^n c_{ij})1}$$
;
- 6
$$x_{gp}^e = \left| \frac{1_{gp}}{c_{gp}} \times w_{gp} - \frac{1_{gp}}{c_{pg}} \times w_{pg} \right|$$
;
- 7
$$w_{gp} = w_{gp} + \left(\frac{1_{gp}}{c_{gp}} \times w_{gp} - \frac{1_{gp}}{c_{pg}} \times w_{pg} \right)$$
;
- 8 end for;
- 9 for all groups Ω_{ij} of BSN(Ω);
- 10 run the Diffusion procedure on Ω_{ij} ;
- 11 end for;
- 12 for all intergroup edges $e: (\langle i, g, p \rangle, \langle 1-i, p, g \rangle)$;
- 13
$$w_{gp} = w_{gp} + \left(\frac{1_{gp}}{c_{gp}} \times w_{gp} - \frac{1_{gp}}{c_{pg}} \times w_{pg} \right)$$
;
- 14 end for;
- 15 for all groups Ω_{ij} of BSN(Ω)
- 16 Return the balancing load vector as w_{ij}^l ;
- 17 end for

4 理论比较

CDE-X 策略在负载均衡过程中的收敛速度只取决于相关因子网络的扩展范式 γ 的大小,而与负载节点的分布无关,与 X 也是无关的. 当 CDE-X 方法执行时, BSN 内各因子网络会并行进行负载均衡过程,误差项 e^k 则是衡量经过 k 步迭代后负载均衡效果的重要指标. 下面的理论比较将说明 CDE-X 比 X 更有优势.

定理 1 Ω 是 BSN(Ω) 的因子网络, 则 $\mu^f(\Omega) < \mu(\text{BSN}(\Omega))$.

证明 BSN(Ω) 可看成是一个连通的完全二分图, 则 Ω 网络图是 BSN(Ω) 网络图的一个真的连通生成支

撑子图.

假设 μ 及 μ^f 分别表示 $\text{BSN}(\Omega)$ 和 Ω 的拉普拉斯特征值, 而 λ 及 λ^f 分别表示 $\text{BSN}(\Omega)$ 和 Ω 网络图的特征值. 将这些特征值从大到小排序, 例如 $\mu_1 > \mu_2 > \dots > \mu_n$; \mathbf{D} 代表图的度对角矩阵, 而 \mathbf{B} 代表图的邻接矩阵, 则有

$$\mu^f(\Omega) = \lambda_1^f(\mathbf{D}(\Omega) + \mathbf{B}(\Omega))$$

而已知, 若 \mathbf{B} 是非负矩阵, 则 \mathbf{B} 的最大特征值随 \mathbf{B} 中元素的增加而递增. 则有:

$$\begin{aligned} \mu^f(\Omega) &= \mu_1^f(\Omega) \\ &= \lambda_1^f(\mathbf{D}(\Omega) + \mathbf{B}(\Omega)) \\ &< \lambda_1(\mathbf{D}(\text{BSN}(\Omega)) + \mathbf{B}(\text{BSN}(\Omega))) \\ &= \mu(\text{BSN}(\Omega)) = \mu_1(\text{BSN}(\Omega)) \end{aligned}$$

证毕.

定理 2 假设 γ 与 γ^f 分别为 $\text{BSN}(\Omega)$ 与 Ω 的扩散矩阵的第二大特征值, 则 $\gamma^f < \gamma$. (根据 $\|e^k\|_2 \leq \gamma^k \|e^0\|_2$, $\gamma^f < \gamma$ 表示若把 CDE-FOS 及 FOS 分别应用在 $\text{BSN}(\Omega)$ 时, 在第 k 步迭代后, CDE-FOS 比 FOS 的误差项 e^k 会有更小些的上界, 从而, CDE-X 的负载均衡效果要更为突出.)

证明 根据文献 [6], 得出 $\gamma = \frac{\lambda_m - \lambda_2}{\lambda_m + \lambda_2}$, $\gamma^f = \frac{\lambda_m^f - \lambda_2^f}{\lambda_m^f + \lambda_2^f}$. 因为 $\text{BSN}(\Omega)$ 可以看成是一个完全两部图 $K_{n,n}$, 因此, 根据图特征定义, λ_2 可看成为 0, 得出 $\gamma^f = \frac{\lambda_m^f - \lambda_2^f}{\lambda_m^f + \lambda_2^f} < 1$; 而 $\gamma = \frac{\lambda_m - \lambda_2}{\lambda_m + \lambda_2} \approx 1$. 因此证得 $\gamma^f < \gamma$.

定理 3 对比 BSN 的谱值计算, CDE-OPT 比 OPT 有更低数量级的计算复杂度.

证明 应用 OPT 算法, $w^k = \left(I - \frac{1}{\lambda^{k+1}}L\right)w^{k-1}$, 最后的迭代误差为 $e^{m-1} = \prod_{i=2}^m \left(I - \frac{1}{\lambda_i}L\right) \sum_{j=2}^m z_j = 0$. OPT 算法不需要额外的变量, 需要知道 $\text{BSN}(\Omega)$ 网络图的不同特征值的个数, 其算法的迭代次数等于拉普拉斯矩阵的特征值个数减 1, 计算特征值的复杂度也达到 $O(n^3)$; 而 CDE-OPT 是只需要知道其因子网络的不同拉普拉斯特征值, 其特征值的计算数量级为 $O(n^2)$. 因此, CDE-OPT 大大降低了计算复杂度.

定理 4 BSN 达到均衡状态前, CDE-OPS 会比 OPS 获得迭代次数 k 的更小些的上界值.

证明 传统的 OPS 算法中, 虽然存在某种情况, 即达到均衡状态时, 网络系统的迭代次数 k 的取值与其图的直径无关, 然而, 针对对称性的 $\text{BSN}(\Omega)$ 网络, 其均衡流量的迭代次数 k 满足下面关系式: $k < \mathbf{D}(\text{BSN}(\Omega))$; 而 CDE-OPS 方案中, 迭代基本发生在因子网络

内部, 且 $d(\Omega) = \frac{1}{2} \mathbf{D}(\text{BSN}(\Omega)) - 1 < \mathbf{D}(\text{BSN}(\Omega))$, 因此, 对于迭代次数 k 的取值, CDE-OPS 将获得 k 的一个更小些的上界值, 从而, CDE-X 降低了迭代次数, 减少了通信流量, 节约了通讯成本.

5 结论

基于 $\text{BSN}(\Omega)$ 一系列的理想特性, 本文提出了一种新的负载均衡策略 CDE-X, 并在理论上对比分析了 CDE-X 的简单、有效性, 论证了基于因子网络的模型 CDE-X 在进行迭代均衡中比标准的单一方案 X 更快的收敛性以及更显著的负载均衡效果, 并且避免了计算 $\text{BSN}(\Omega)$ 的 $2n^2 \times 2n^2$ 矩阵的拉普拉斯特征值, 而只需计算因子网络的拉普拉斯特征值, 大大降低了达到均衡迭代状态的负载迁移的流量.

虽然 CDE-X 在迭代过程中发生在网络链接上的迁移负载不是 l_2 -minimal, 但是, 可以考虑如何更能降低部间链接的通信流量, 使得任务图中通信密集的子任务尽可能映射到同一个因子网络中, 这也是今后对于该领域的一个重点研究方向. 为体现 BSN 的实际应用, 还需实验性研究负载均衡及任务划分等重要问题.

参考文献

- [1] 蒋江, 张民选, 廖湘科. 基于多种资源的负载均衡算法的研究[J]. 电子学报, 2002, 30(8): 1148 - 1152.
Jiang Jiang, Zhang Min-xuan, Liao Xiang-ke. Study on load balancing algorithms based on multiple resources [J]. Acta Electronica Sinica, 2002, 30(8): 1148-1152. (in Chinese)
- [2] 杨际祥, 谭国真, 王荣生. 并行与分布式计算动态负载均衡策略综述[J]. 电子学报, 2010, 38(5): 1122 - 1130.
Yang Ji-xiang, Tan Guo-zhen, Wang Rong-sheng. A survey of dynamic load balancing strategies for parallel and distributed computing [J]. Acta Electronica Sinica, 2010, 38(5): 1122 - 1130. (in Chinese)
- [3] Cybenko George. Dynamic load balancing for distributed memory multiprocessors [J]. Journal of Parallel and Distributed Computing, 1989, 7(2): 279 - 301.
- [4] Muthukrishnan S, Ghosh Bhaskar, Schultz Martin H. First and second order diffusive methods for rapid, coarse, distributed load balancing [A]. Proceedings of the Eighth Annual ACM Symposium on Parallel Algorithms and Architectures, 1996 [C]. USA: Springer-Verlag, 1996.
- [5] Ralf Diekmann, Andreas Frommer, Burkhard Monien. Efficient schemes for nearest neighbor load balancing [J]. Parallel Computing, 1999, 25(7): 789 - 812.
- [6] Robert Elsasser, Andreas Frommer, Burkhard Monien, et al. Optimal and alternating-direction load balancing schemes [A]. Proceedings of Euro-Par' 99, 1999 [C]. Berlin: Springer-Ver-

- lag, 1999.
- [7] Robert Elsasser, Burkhard Monien, Robert Preis. Diffusive load balancing schemes on heterogeneous networks[A]. Proceedings of twelfth ACM Symposium on Parallel Algorithms and Architectures, 2000[C]. New York: ACM, 2000.
- [8] Rotaru Tiberiu, Negeli Hans-Heinrich. Dynamic load balancing by diffusion in heterogeneous systems[J]. Journal of Parallel and Distributed Computing, 2004, 64(4): 481 - 97.
- [9] 杨际祥, 谭国真, 王凡, 等. 一种大规模分布式负载均衡策略[J]. 电子学报, 2012, 40(11): 2226 - 2231.
Yang, Ji-xiang, Tan Guo-zhen, Wang Fan, et al. A load balancing strategy for large-scale distributed computing [J]. Acta Electronica Sinica, 2012, 40(11): 2226 - 2231. (in Chinese)
- [10] Parhami Behrooz. Swapped interconnection networks: Topological, performance, and robustness attributes[J]. Journal of Parallel and Distributed Computing, 2005, 65 (11): 1443 - 1452.
- [11] Chen Wei-dong, Xiao Wen-jun, Parhami Behrooz. Swapped (OTIS) networks built of connected basis networks are maximally fault tolerant [J]. IEEE Transactions on Parallel and Distributed Systems, 2009, 20(3): 361 - 366.
- [12] 陈卫东, 肖文俊. Biswapped 网络(BSN)的拓扑性质研究: 点对称性和极大容错性[J]. 计算机学报, 2010, 33(5): 822 - 832.
Chen Wei-dong, Xiao Wen-jun. Topological properties of biswapped networks (BSNs): node symmetry and maximal fault tolerance [J]. Chinese Journal of Computers, 2010. 33 (5): 822 - 832. (in Chinese)
- [13] Xiao Wen-jun, Chen Wei-dong, He Ming-xin, et al. Biswapped networks and their topological properties [A].

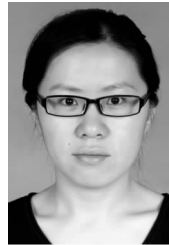
Proceedings of Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007[C]. Qingdao: Inst of Elec and Elec Eng Computer Society, 2007.

- [14] Arndt, H. Load balancing: dimension exchange on product graphs[A]. Proceedings of eighteenth International Parallel and Distributed Processing Symposium, 2004[C]. Los Alamitos: IEEE Comput Soc, 2004.
- [15] Holger Arndt. On finite dimension exchange algorithms[J]. Linear Algebra and Its Applications, 2004, 380: 73 - 93.

作者简介



童朝南 男, 1954 年 6 月出生, 安徽无为
人. 北京科技大学自动化学院教授、博士生导师.
主要研究方向为集群优化、生产过程自动化、生
产过程智能控制、鲁棒控制理论及应用.
E-mail: tcn@ies.ustb.edu.cn



孙丽婷(通信作者) 女, 1985 年 4 月出生,
山东邹城人. 北京科技大学博士研究生. 主要研
究方向为集群优化、并行计算及负载均衡.
E-mail: sun-liting@163.com