

文章编号: 2096-1472(2017)-01-12-03

一种基于Torus网络的高效随机Oblivious路由算法

任一曼

(天津工业大学计算机科学与软件学院, 天津 300387)

摘要: 一个好的路由算法应同时满足: 最小的路由跳数以减小传输延时, 保持通讯的局域性; 最大的平均情况和最坏情况吞吐率; 简单的路由器结构。随机Oblivious路由算法在低功耗并行计算机互连网络以及片上网络中得到广泛应用。针对Torus网络下已提出的Oblivious路由算法所需虚通道数目多的缺点, 提出了随机Oblivious路由算法WRD, 该算法仅使用两条虚拟通道即可实现算法的无死锁性。通过仿真对所提算法的性能进行了验证, 结果表明, 该算法与使用两条虚拟通道的OITURN路由算法相比, WRD路由算法在所有通讯模式下的网络吞吐率均有所提升。与使用四条虚拟通道的RLB算法相比, 新提出的WRD路由算法性能接近于RLB算法, 甚至在多个通讯模式下的网络吞吐率要好于RLB算法, 而且WRD路由算法仅使用两条虚拟通道, 降低了网络系统成本和功耗。

关键词: Torus网络; 随机Oblivious路由算法; 平均情况网络吞吐率; 最坏情况网络吞吐率; 虚拟通道

中图分类号: TP393

文献标识码: A

An Efficient Random Oblivious Routing Algorithm Based on the Torus Network

REN Yiman

(School of Computer Science and Software, Tianjin University of Technology, Tianjin 300387, China)

Abstract: A good routing algorithm should satisfy: the minimum number of routing hops to reduce the propagation delay and maintain the localization of the communication. The maximum average case and the worst case throughput. The simple router structure. The random Oblivious routing algorithm is widely used in low-power parallel computer interconnection networks and on-chip networks with its high routing flexibility. Aiming at the shortcomings of the number of virtual channels required by the Oblivious routing algorithm proposed in the Torus network, the paper proposes a random Oblivious routing algorithm WRD. The algorithm uses only two virtual channels to achieve the deadlock-freeness of the algorithm. The performance of the proposed algorithm is verified through simulation. The results show that the algorithm has improved the network throughput of WRD routing algorithm in all communication modes compared with the OITURN routing algorithm using two virtual channels. Compared with the RLB algorithm using four virtual channels, the newly proposed WRD routing algorithm is close to the RLB algorithm, and even in multiple communication modes, the network throughput is better than that of the RLB algorithm, and the WRD routing algorithm uses only two virtual channels, reducing network system costs and power consumption.

Keywords: the Torus network; the random oblivious routing algorithm; average network throughput; worst case network throughput; virtual channels

1 引言(Introduction)

Torus网络, 也叫k元n立方网络, 是一类非常重要的互联网络, 在所有的应用领域中占据主导地位^[1,2]。Torus网络实际上可以理解成带有环绕通道的mesh网络, 在Torus网络中设计一种无死锁的高效路由算法具有很重要的意义^[3]。

网络吞吐量和传输延时在设计路由算法时是最重要的性能指标。理想情况下, 我们设计一个路由算法, 希望满足最大的最坏情况和平均情况网络吞吐率, 以及最小的传输延时。虚通道将一条物理通道分为多条逻辑通道并多路复用该物理通道^[4]。我们通过增加虚拟通道的数目, 并对虚通道的使用做一定的路由限制, 来解决网络中的死锁问题。虚拟通道

的数目并不是越多越好, 研究使用虚拟通道少的无死锁路由算法具有重要意义。

为了进一步提高最坏情况和平均情况网络吞吐率, 本文提出了一种新的高吞吐率随机Oblivious路由算法——WRD(Weighted Random Double Direction)算法。

2 网络吞吐量分析(Network throughput analysis)

网络容量被定义为在均匀随机的通讯模式下, 网络中可持续的最大吞吐率。网络容量具体可以使用(消息数或微片数/节点/周期)^[5], 即每个节点每个周期产生的消息或微片数量来表示。普遍的看法是网络的等分面是网络负载最重的地方, 假设一个k*k的2D-Mesh网络, 其中k为偶数, 等分面将

这一网络划分为两部分，每部分都有 $k \cdot \frac{k}{2}$ 个节点的两部分，这两部分各自有 k 个通道连接到另外一部分。假设在某一时刻，网络达到饱和，从其中一部分产生的消息在均匀随机通讯模式下，有一半消息会穿过等分面到达另一部分^[6]，则网络容量 N_C 可以由以下等式得出：

$$N_C \cdot \left(k \cdot \frac{k}{2} \right) \cdot \left(\left(k \cdot \frac{k}{2} \right) \div k^2 \right) = k \quad (1)$$

则当 k 为偶数时， $N_C = \frac{4}{k}$

同理可得，当 k 为奇数时， $N_C = \frac{4k}{k^2 - 1}$

已知一个网络有 N 个节点，其网络通讯矩阵 $\Lambda = (\lambda_{i,j})$ 是一个 $N \times N$ 的矩阵，其中 $\lambda_{i,j}$ 表示从源节点 i 到目的节点 j 的通讯流量。如果有下列数学式成立， Λ 是双随机和可接受的矩阵。

已知一个随机Oblivious路由算法 R 和一个双随机、可接受的通讯矩阵 Λ ，通道负载 $\gamma_{\max}(R, \Lambda)$ 是所有通道中通讯量期望值最高的通道的负载。而最坏情况下的通道负载 $\gamma_{wc}(R, \Lambda)$ 则可以表示为： $\gamma_{wc}(R, \Lambda) = \max_{\Lambda \in P} (\gamma_{\max}(R, \Lambda))$ ，其中 P 为所有双随机、可接受的通讯矩阵的集合。

如果采用穷举法求解这一问题，则算法的时间复杂度为 $O(CN!)$ ，其中 C 为通道数， N 为节点数。这一问题实际上属于最大权匹配问题^[7]，借助于匈牙利算法可以将求解这一问题的时间复杂度降为 $O(CN^3)$ ，因此随机Oblivious路由算法 R 最坏情况网络吞吐率 $TP_{wc}(R)$ 的求解公式为：

$$TP_{wc}(R) = b_c \div \gamma_{wc}(R, \Lambda) = (\gamma_{wc}(R, \Lambda))^{-1} \quad (2)$$

除了特别声明，本文将使用 $\Theta_{wc}(R)$ 表示随机Oblivious路由算法 R 最坏情况网络吞吐率与网络容量之比：

$$\Theta_{wc}(R) = \frac{TP_{wc}(R)}{N_C} \quad (3)$$

随机Oblivious路由算法 R 平均情况下的网络吞吐率 $\Theta_{avg}(R)$ 是在一定的时间周期 T 中，对随机产生的通讯矩阵所导致的网络负载 $\lambda_{\max}(R, \Lambda)$ 求平均值得出：

$$\Theta_{avg}(R) = \left(\frac{1}{T} \sum_T (\gamma_{\max}(R, \Lambda)) \right) / N_C \quad (4)$$

本文中 $|T| = 100000$ 。

3 一种新的Torus网络随机Oblivious路由算法(A new random oblivious routing algorithm for torus networks)

本节我们首先介绍一种新的随机Oblivious路由算法WRD算法，然后对其性能和虚拟通道的使用情况进行分析，最后

对其无死锁性进行证明。

3.1 提出WRD算法的过程

通过RLB算法，我们可以得到两点启示：一是其核心思想的运用。RLB算法允许数据包沿较长路径和较短路径两个方向双向路由，只是沿较长路径路由的概率比较短路径路由的概率小一些。我们借鉴这种思想及OITURN路由算法的思想，提出了新的路由算法WRD(Weighted Random Double Direction)，新的路由算法允许消息以一定的概率沿较长路径和较短路径两个方向由源节点传输至目的节点，这很好的均衡了整个网络的负载，从而提高了网路性能。同时，减少了虚拟通道的数目，从而降低了系统成本和能耗，并缩短了源节点到目的节点的平均距离，从而减小延迟。

3.2 WRD算法的设计思想

本文将RLB的核心思想以及OITURN路由算法的思想应用到Torus网络下的Oblivious路由算法中提出来了WRD路由算法。Torus网络可以看作是带回绕的Mesh网络，在不考虑回绕的情况下，对于任何给定的源节点和目的节点对之间，最多存在两条最小的，仅有一个转弯的路径。这两条路径在维度遍历(XY或YX)的顺序上不同，WRD算法随机的选择第一个遍历的维度。可以将WRD算法理解为ROMM路由的受限版本，其中间节点为最小矩形的四个角中除源节点目的节点，另外两个节点中的一个。数据包由源节点路由至中间节点不再是只能走最短路径，而是允许消息沿较长路径和较短路径双向路由，但分配给沿较长路径路由的概率要小于沿较短路径路由的概率。

3.3 算法描述

根据上节所述算法的设计思想，具体算法描述如下：

在 $K \times K$ 的Torus网络中，假定源节点 S 的坐标为 (S_1, S_2) ，目的节点 D 的坐标为 (D_1, D_2) ，我们生成一个距离向量 $\Delta = (\Delta_1, \Delta_2)$ ，其中 $\Delta_1 = \min(|S_1 - D_1|, K - |S_1 - D_1|)$ 。我们计算出在某一维上，沿较长路径路由的概率为 $P_1 = (K - \Delta_1)/K$ ，沿较短路径路由的概率为 $P_s = \Delta_1/K$ 。

(1)对于给定的源节点—目的节点对，首先随机的选择第一个遍历的维度，即选择X维或Y维的概率各占50%。一旦选择了第一个遍历的维度，就确定了中间节点。

(2)由源节点到中间节点既可以沿较长路径路由，也可以沿较短路径路由。沿较长路径路由的概率是 $(K - \Delta_1)/K$ ，沿较短路径路由的概率是 Δ_1/K 。

(3)转入下一个维度,将消息由中间节点路由至目的节点,同样允许消息沿较长路径和较短路径双向路由,沿较长路径路由的概率是 $(K-\Delta_i)/K$,沿较短路径路由的概率是 Δ_i/K 。

4 仿真与分析(Simulation and analysis)

本章将根据不同的网络规模在不同的通讯模式下对新提出来的WRD路由算法的性能进行仿真,并与现有的DOR算法以及几种Oblivious路由算法进行仿真比较,如ROMM、VAL、OITURN、RLB。仿真实验使用的交换技术是虫孔交换技术,每条消息长度设为16位,缓冲池大小设为64位,通讯模式如表1所示,其中DOR-WC表示使DOR路由算法吞吐量为最差的一种通讯模式。

表1 验仿真的几种通讯模式

Tab.1 Several communication modes of simulation	
转置	由节点(x,y)传输至(y,x)
互补	由节点(x,y)传输至(k-x-1,k-y-1)
龙卷风	由节点(x,y)传输至(x+k/2-1,y)
DOR-WC	由节点(x,y)传输至(k-y-1,k-x-1)
平均情况	对随机产生的100000个通讯矩阵的网络吞吐率求平均值
最坏情况	最坏情况通讯矩阵导致的最差网络吞吐率

表2是对于k*k的2D-Torus,当k为奇数时,WRD路由算法和DOR、ROMM、VAL、OITURN、RLB在所有通讯模式下的网络吞吐率。由表可以看出,当K为奇数时。

(1)与OITURN相比较,本文新提出的WRD路由算法平均情况下的网络吞吐率有了一定的提高,其他几种通讯模式下的网络吞吐率也与OITURN持平。在平均情况下,当网络基数K=9时,DOR、ROMM、VAL、OITURN的吞吐率分别为0.308,0.336,0.5,0.371,WRD路由算法的吞吐率为0.435,均高于DOR、ROMM、OITURN,虽低于VAL,但在其他几种通讯模式下,WRD路由算法的表现都优于VAL。

(2)与RLB相比较,新提出的WRD路由算法在互补、DOR-WC的通讯模式下的吞吐率要高于RLB,平均情况、最坏情况下的吞吐率虽不及RLB但相差不大,且WRD仍具有两方面的优势,一是新提出的WRD路由算法只需要两条虚拟通道,而RLB算法需要四条虚拟通道,这一点有非常重要的意义,之前我们已经提过了,虚拟通道的数目并不是越多越好,增加虚拟通道的数目反而会增加成本与功耗,这是一个很大的改进。二是新提出的路由算法WRD缩短了源节点到目的节点的平均距离,从而降低了延时。

表2 当k为奇数时网络吞吐率

Tab.2 Network throughput when k is odd						
9*9 Torus网络	DOR	ROMM	VAL	OITURN	RLB	WRD
转置	0.278	0.556	0.5	0.556	0.687	0.556
互补	0.556	0.362	0.5	0.556	0.476	0.556
DOR-WC	0.278	0.255	0.5	0.556	0.529	0.556
平均情况	0.308	0.336	0.5	0.371	0.597	0.435
最坏情况	0.278	0.173	0.5	0.278	0.349	0.278

图1所示为网络规模k为奇数时WRD路由算法和DOR、VAL、OITURN、RLB在不同网络规模下的平均情况下网络吞吐率的比较结果,其中横坐标为网络规模k选择奇数3、5、7、9;纵坐标为平均情况下网络吞吐率。由表可知,在平均情况下,新提出的WRD路由算法的吞吐率要高于DOR和OITURN路由算法,并且随着网络规模的增大,吞吐量越来越大。虽然WRD算法在平均情况下的吞吐率不及RLB算法,但WRD路由算法只用到了两条虚拟通道,RLB算法用到了四条虚拟通道,这是很大的改进。

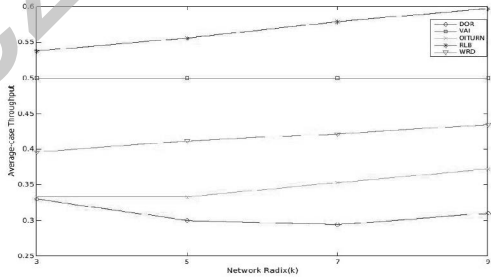


图1 为奇数时平均情况网络吞吐率

Fig.1 Average time network throughput when k is odd

5 结论(Conclusion)

二维Torus网络上的WRD路由算法借鉴了RLB算法和OITURN算法的核心思想,允许数据包按一定概率沿较长和较短路径双向路由,实现了由源节点到目的节点多条路径路由的Oblivious路由算法,该算法只使用了两条虚拟通道,与OITURN路由算法相比,新提出的WRD路由算法在所有通讯模式下的网络吞吐率较OITURN路由算法相比均有所提升。与使用四条虚拟通道的RLB算法相比,新提出的WRD路由算法性能接近于RLB算法,甚至在多个通讯模式下的网络吞吐率要好于RLB算法,而且WRD路由算法仅使用两条虚拟通道,降低了网络系统成本和功耗。增加虚拟通道的数目时,几个消息的微片会复用同一条物理通道,减慢了消息的传播速度,增加了通道的传播延迟,而且还会增加可选择的路由

(下转第8页)