

基于 3D-Mesh 网络的一种 Oblivious 路由算法

王振江, 段新明

(天津工业大学 计算机科学与软件学院, 天津 300387)

摘 要: 3D-Mesh 就是三维片上网络中一种常见拓扑结构. 本文针对 3D-Mesh 网络提出了一种新的高吞吐率的随机 Oblivious 路由算法(Double Plane Turn, DPT). DPT 算法通过降维, 将 3D-mesh 网络转化为两个互相垂直 2D-mesh, 然后在 2D-mesh 网络下使用 O1Turn 路由算法. 由于 O1Turn 路由算法具有良好的平均情况下的网络吞吐率和不错的延时, 通过仿真实验证明, 和以往的路由算法在平均情况下的网络吞吐率相比, 本算法都实现了较好的结果.

关键词: 3D-Mesh; Oblivious 算法; 平均情况网络吞吐率; DPT

中图分类号: TP393

文献标识码: A

文章编号: 1000—7180(2018)03—0023—05

An Oblivious Routing Algorithm Based on 3D-Mesh Network

WANG Zhen-jiang, DUAN Xin-ming

(College of Computer Science and Software, Tianjin Polytechnic University, Tianjin 300387, China)

Abstract: 3D-Mesh is a common topology in a three-dimensional on-chip network. In this paper, a new high-throughput random Oblivious routing algorithm (Double Plane Turn, DPT) is proposed for 3D-Mesh network. The DPT algorithm transforms the 3D-mesh network into two mutually perpendicular 2D-meshes by reducing the dimension, and then uses the O1Turn routing algorithm on the 2D-mesh network. As the O1Turn routing algorithm has a good average network throughput and good delay, simulation experiments show that compared with the previous routing algorithm in the average case of network throughput, the algorithm has achieved good results.

Key words: 3D-Mesh; Oblivious routing; average-case throughput; DPT

1 引言

随着三维集成电路的快速发展^[1-2], 三维片上网络也受到了广泛关注与发展, 从而解决了系统集成度增加导致的通信瓶颈问题, 进一步提高 Soc 性能^[3]. 因此 3DNoc 将成为 Noc 的发展趋势和主要研究方向.

在片上系统中, 路由算法是影响网络性能的一个重要因素. 根据常见的路由算法分类方法, 可以将路由算法分为确定性路由和自适应路由^[4]. 一个好的路由算法, 应该具有较好的网络吞吐率和低传输延时. 确定性路由算法虽然简单, 但路径单一, 容易造成网络资源浪费, 自适应路由算法虽然可以自

动避开网络拥塞节点, 但实现复杂且容易死锁, 例如 DOR 路由算法^[5]. Oblivious 路由算法则是一种在进行路由决策时不考虑网络状态的路由算法, 且同时可以随机的选择路由路径, 实现了较高的路由灵活性. 常见的 3D-mesh 下的 Oblivious 路由算法有从 2D-mesh 扩展到 3D-mesh 的 VAL、ROMM、O1TURN 等算法, 此外还有 RPM、USM 等算法. VAL 路由算法^[6]虽然在最坏情况下有不错的网络吞吐率, 但平均情况下网络吞吐率不是特别好, 而且其路由延时比较高. ROMM 路由算法^[7]具有良好的平均情况下网络吞吐率, 但其最坏情况下网络吞吐率表现非常糟糕. O1TURN 路由算法^[8]在 2Dmesh 网络中很好地均衡了最坏情况和平均情况

网络吞吐率以及消息传输延时,但在 3D-mesh 网络中的吞吐率都不是特别好. RPM 算法^[9]在 mesh 网络阶数为偶数时实现了最优的最坏情况网络吞吐率,且具有良好的平均情况网络吞吐率. USM 路由算法^[10]在 mesh 网络阶数为基数时实现了最优的最坏情况下的网络吞吐率,但传输延时较高.

衡量网络性能的两个主要因素是网络的吞吐率和传输延时. 为了进一步获得更优的网络吞吐率和降低传输延时,本文针对三维片上网络下的 Mesh 拓扑结构设计了一种新的 Oblivious 随机路由算法, DPT 算法. 该算法实现了更好的网络负载,同时相较于 RPM 算法,进一步提高了平均情况下的网络吞吐率.

2 相关工作

在三维片上网络系统中,影响片上系统性能的因素主要有拓扑结构、路由算法、子网划分及无线片上网络结构等. 其中,拓扑结构和路由算法则是影响片上系统性能的主要因素. 拓扑结构保证了每个节点可以发送数据到其它节点,路由算法则决定数据包从源节点到目的节点的路由路径.

2.1 拓扑结构

传统的拓扑结构主要包括二维上的 Mesh 结构和 Torus 结构,以及三维上的 Mesh 结构和胖树结构. 其中,2D-Mesh 结构如图 1 所示.

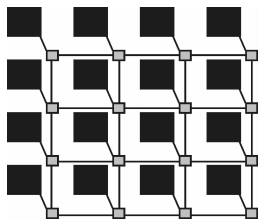


图 1 4×4 的 2D-Mesh 架构

3D-mesh 的结构在每层上仍采用 2D-Mesh 结构,然后实现层层相连. 3D-mesh 的结构如图 2 所示.

2.2 网络吞吐率

最坏情况和平均情况下的网络吞吐率是常用的一种评价一个网络性能的重要指标. 其中,利用网络容量去计算 Oblivious 路由算法在最坏情况下的网络吞吐率. 平均情况下的网络吞吐率使用了另一种方法去做评估. 不考虑流量控制的情况,假设每个节点在一个时钟周期内向邻居节点发送一个微片或者消息,即网络带宽 $b_c=1$.

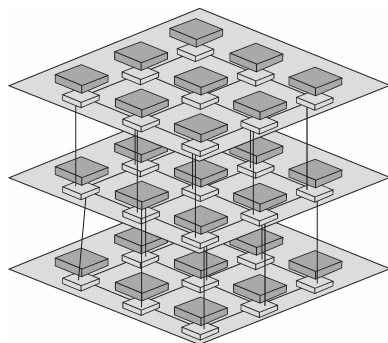


图 2 3D-mesh 结构(3×3×3)

(1)网络容量. 当一个网络在均匀随机的通讯模式下,网络中可持续的最大通道负载^[8]. 因为所有的通道带宽被假定为 1,那么网络容量意味着网络带宽除以该网络的最大通道负载,即网络容量为最大通道负载 γ^* 的倒数. 对于任一 k 维 n 个节点的 mesh 网络有

$$\gamma^* = \begin{cases} \frac{k}{4}, & k \text{ is even} \\ \frac{k^2 - 1}{4k}, & k \text{ is odd} \end{cases} \quad (1)$$

(2)最坏情况下网络吞吐率. 已知一个随机 Oblivious 路由算法 R 和一个双随机、可接受通讯矩阵 $\mathbf{A}$,最大通道负载 $\gamma(R, \mathbf{A})$ 表示随机 Oblivious 路由算法 R 作用在通讯矩阵 $\mathbf{A}$ 中所有通道中负载最高的通道的负载^[9].

通常,最坏情况下的网络吞吐率等于最坏情况下最大通道负载 $\gamma(R, \mathbf{A})$ 的倒数. 但我们定义标准化的最坏情况下的网络吞吐率为最坏情况下的网络吞吐率与网络容量之比,记作 $\Theta_{wc}(R)$:

$$\Theta_{wc}(R) = \left(\frac{\gamma_{wc}(R)}{\gamma^*} \right)^{-1} \quad (2)$$

除非特此声明,在下文中我们都将用 $\Theta_{wc}(R)$ 来表示最坏情况网络吞吐率.

(3)平均情况下网络吞吐率. 一个路由算法的平均情况下的网络吞吐率是通过计算在一定的时钟周期 T 内随机产生的通讯矩阵的最大网络负载求平均值所得. 平均情况下的网络吞吐率,记作 $\Theta_{avg}(R)$, 其中 $|T| = 100\ 000$ ^[9].

$$\Theta_{avg}(R) = \frac{1}{|T|} \sum_{\mathbf{A} \in T} \left(\frac{\gamma(R, \mathbf{A})}{\gamma^*} \right)^{-1} \quad (3)$$

3 DPT 路由算法

本文提出了一种新的在 3D-mesh 网络下的随机路由算法, Double Plane Turn 算法,简称 DPT

算法.

3.1 算法描述

DPT 算法将 3D-mesh 网络降维到两个 2Dmesh 网络,即 XY 平面和 YZ 平面.与 RPM 算法不同,DPT 算法将数据包从源节点路由至目的节点只需要两个阶段.第一个阶段,数据包在 XY 平面内使用 01Turn 算法路由至一个中间节点,第二阶段数据包继续从中间节点在 YZ 平面内继续使用 01Turn 算法路由至目的节点.在 3D-mesh 网络中,一个节点必然可以通过一个水平平面和相垂直的垂直平面到达目的节点.而每个平面都相当于一个 2D-Mesh,然后使用 2D-Mesh 下的 01Turn 算法进行路由.由于 01Turn 路由算法具有良好的平均情况下的网络吞吐率和不错的延时,所以 DPT 算法在降维后的 2D-mesh 网络上使用了 01Turn 算法. DPT 算法总共包含 4 种可能的路径,图 3 给出了 4 种可能的路由路径,分别是 XY-YZ、XY-ZY、YX-YZ 和 YX-ZY.

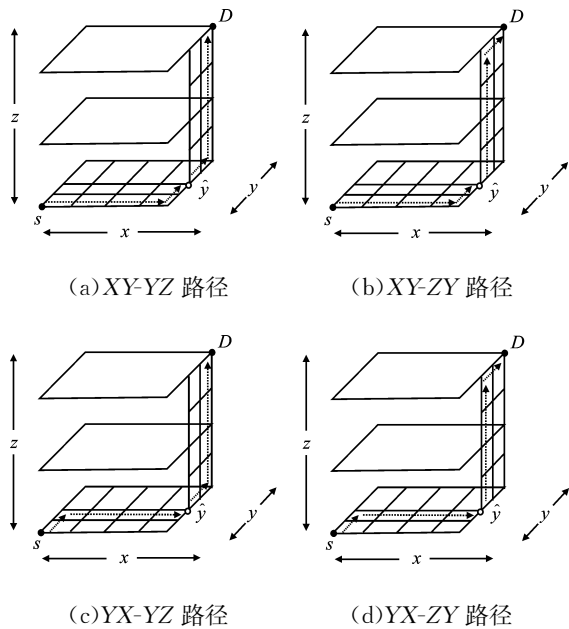


图3 DPT 算法四种可能的路径

假设源节点坐标 s 为 (x_0, y_0, z_0) , 目的节点 d 为 (x_1, y_1, z_1) 则 $\hat{y}$ 表示在 Y 轴上随机取的一个点, 则 DPT 算法路由算法可描述为:

- (1) 随机选择一个 $\hat{y}$, 其中 $\hat{y} \in (0, k-1)$;
- (2) 由源节点 (x_0, y_0, z_0) 和节点 $(x_1, \hat{y}, z_0)$ 确定 XY 平面, 数据包使用 2DMesh 网络下的 01Turn 算法经过 XY 平面到达中间节点 $(x_1, \hat{y}, z_0)$;
- (3) 由中间节点 $(x_1, \hat{y}, z_0)$ 和目的节点 (x_1, y_1, z_1) ,

z_1) 确定 YZ 平面, 数据包再次使用 01Turn 算法经过 YZ 平面到达目的节点 (x_1, y_1, z_1) .

声明 1 DPT 路由算法为无死锁路由算法.

证明: 在 3D-mesh 中通过降维将 3D-mesh 转化为两个互相垂直且只有一个相交点的 2D-mesh, 根据路由向前性, DPT 算法将数据包经由水平平面 XY 和垂直平面 YZ 使用了两次 2DMesh 下的 01Turn 算法最终到达目的节点. 由于在 2D-Mesh 下的 01Turn 算法是无死锁路由算法, 所以 DPT 也是无死锁路由算法.

3.2 网络吞吐率分析

本节主要是分析 DPT 算法在最坏情况下和平均情况下的网络吞吐率.

最坏情况下和平均情况下的网络吞吐率是衡量网络性能的一个常用指标^[2]. 一个好的路由算法, 应该具备不错的网络吞吐率和较低的传输延时两个特征. 由第二节可知, 计算最坏情况下的网络吞吐率首先要计算当前网络的最大负载或者网络吞吐率(等于最大负载的倒数).

声明 1 在 $K * K * K$ 的 3D-mesh 网络中使用 DPT 算法的通道最大负载量为 $K/2$.

证明: DPT 算法的实现是通过将 3D-mesh 网络转化为两个 2DMesh 网络(互相垂直, 避免了死锁)然后进行无死锁路由. 2DMesh 网络中已经证明最短路径路由 XY 路由和 YX 路由有等概率的最大通道负载, 最大通道负载为 $k/2$. DPT 算法虽然经过两次转换, 但都是相当于在 2DMesh 网络下的 01Turn 路由^[3]. 所以, DPT 算法的最大通道负载也为 $K/2$.

声明 2 DPT 算法的实现了最大的最坏情况下的网络吞吐率, 当 k 为偶数时, 最大最坏情况下的网络吞吐率为 0.5; 当 k 为奇数时, 最大最坏情况下的网络吞吐率为 $0.5 * (1 - (1/k * k))$.

证明: 由声明 1, 可知 $\gamma_{wc}(DPT) = k/2$. 由第二节可知, 当 k 为偶数时, $\gamma^* = k/4$, 因此

$$\Theta_{wc}(DPT) = \left\lceil \frac{\frac{k}{2}}{\frac{k}{4}} \right\rceil^{-1} = 0.5 \quad (4)$$

当 k 为奇数时, $\gamma^* = (k^2 - 1)/4k$, 因此

$$\Theta_{wc}(DPT) = \left\lceil \frac{\frac{k}{2}}{\frac{(k^2 - 1)}{4k}} \right\rceil^{-1} = \left(\frac{k}{2} \cdot \frac{4k}{(k^2 - 1)} \right)^{-1} = \frac{k^2 - 1}{2k^2} = 0.5 \left(1 - \frac{1}{k^2} \right) \quad (5)$$

3.3 平均传输延时分析

算法的平均传输延时通常可以根据路由的平均跳步数计算得到. 在计算平均跳步数时, 要对所有的起始节点到目的节点的跳步数求平均值. 对于一个 $k \times k \times k$ 阶的 3D-mesh 网络, DOR 算法在 3D-mesh 结构中的平均跳步数为 $3(k^2-1)/(3k)$. 其中, 每一维上的平均跳步数为 $(k^2-1)/3k$.

RPM 算法需要在 Z 轴上传输两次, 所以 Z 轴上的平均跳步为 $2(k^2-1)/(3k)$, 2D-mesh 下 01Turn 的传输延时也为 $2(k^2-1)/(3k)$, 因此 RPM 算法的平均传输延时为 $4(k^2-1)/(3k)$. USM 算法则需要五步, X、Y 轴分别传输两次, 故 X、Y 轴上的平均跳步数为 $2(k^2-1)/(3k)$, 所以平均传输延时为 $5(k^2-1)/(3k)$. DPT 路由算法的过程总共需要四步, 所以平均传输延时为 $4(k^2-1)/3k$. 虽然和 RPM 路由算法的平均传输延时相同, 但实现了更好的平均情况下的网络吞吐率. 与 USM 算法相比则实现了更短的传输延时.

4 仿真与分析

本节主要是使用常见的几种路由算法以及 DPT 路由算法在几种不同通讯模式下进行仿真实验, 然后对比各个算法的数据. 仿真实验对比的算法如表 1 所示.

表 1 路由算法描述

算法	算法描述
DOR	先沿着 x 轴路由, 接着 y 轴, 最后 z 轴
VAL	随机选择一个中间节点, 使用 DOR 算法先路由至此节点, 接着再由此节点使用 DOR 算法路由至目的节点
ROMM	可以称之为有限制的 VAL 算法, 中间节点在由起始节点和目的节点形成的最小子立方体内选择
01Turn	随机选择 XYZ, XZY, YXZ, YZX, ZXY, ZYX 方向上的维序路由, 各概率为 1/6
RPM	在垂直方向上(Z 轴)随机路由至一个点, 然后在 XY 面上使用 01Turn 路由, 接着在 Z 轴路由至目的节点

仿真实验的通讯模式如表 2 所示.

表 2 通讯模式

最坏情况	最坏情况通讯矩阵导致的最差网络吞吐率
平均情况	对随机产生的 10^5 个通讯矩阵的网络吞吐率求平均值
转置	由源节点(x, y, z)路由至目的节点(y, z, x)
互补	数据包由节点(x, y, z)路由至节点($k-x-1, k-y-1, k-z-1$)
DOR-WC	数据包由源节点(x, y, z)路由至目的节点($k-z-1, k-y-1, k-x-1$)

通过仿真实验, DPT 算法在网络阶数为 $4 \times 4 \times$

4 和 $6 \times 6 \times 6$ 的 3D-mesh 网络下, 平均情况下的网络吞吐率远优于其它算法. 与 VAL、DOR、ROMM、01Turn、RPM 算法相比, 分别平均提高了 7%, 74%, 31%, 12%, 8%. 在最坏情况下, DPT 算法和 RPM 算法一样实现了最优的网络吞吐率, 在网络阶数为偶数的情况下达到了网络容量的二分之一, 远优于其它算法.

表 3 实验结果

$3 \times 3 \times 3$	VAL	DOR	ROMM	01Turn	RPM	DPT
最坏情况	0.5	0.22	0.185	0.21	0.44	0.44
平均情况	0.5	0.3	0.37	0.413	0.49	0.523
转置	0.5	0.22	0.30	0.44	0.5	0.52
互补	0.5	0.67	0.3	0.5	0.5	0.5
DOR-WC	0.5	0.22	0.33	0.33	0.5	0.5
随机	0.5	1	0.912	1	0.5	0.5
$4 \times 4 \times 4$	VAL	DOR	ROMM	01Turn	RPM	DPT
最坏情况	0.5	0.125	0.203	0.25	0.5	0.5
平均情况	0.5	0.315	0.424	0.471	0.5	0.537
转置	0.5	0.25	0.327	0.5	0.50	0.54
互补	0.5	0.5	0.308	0.5	0.5	0.5
DORWC	0.5	0.125	0.214	0.25	0.5	0.5
随机	0.5	1	0.811	1	0.5	0.5
$6 \times 6 \times 6$	VAL	DOR	ROMM	01TU-RN	RPM	DPT
最坏情况	0.5	0.083	0.126	0.187	0.5	0.5
平均情况	0.5	0.308	0.435	0.488	0.5	0.543
转置	0.5	0.25	0.298	0.5	0.5023	0.54
互补	0.5	0.5	0.23	0.5	0.5	0.5
DORWC	0.5	0.083	0.171	0.187	0.5	0.5
随机	0.5	1	0.742	1	0.5	0.5

5 结束语

本文提出了一种 3D-Mesh 网络下的随机 Oblivious 路由算法, DPT 算法. 与以往的几种 Oblivious 路由算法相比, 无论是在网络吞吐率和传输延时下, 都有不错的改进. 相对于 RPM 算法, 在传输延时相同的情况西下 DPT 算法实现了平均情况下更好的网络吞吐率, 且路径选择更加自由. 相较于 USM 算法, DPT 算法实现了更短的传输延时, 且只需要 2 个虚拟通道就可以避免死锁, 而 USM 算法则需要 3 个虚拟通道才可以实现避免死锁. 此外, RPM 算法从源节点到目的节点需要 3 次转弯, USM 算法则需要四次转弯, 而 DPT 算法最少需要 2 次、最多 3 次转弯就可以到达目的节点.

参考文献:

- [1] 刘招山. 片上网络若干关键技术的研究[D]. 厦门:厦门大学, 2014.
- [2] 张大坤, 黄翠, 宋国治. 三维片上网络研究综述[J]. 软件学报, 2016, 26(1): 155-187.
- [3] 李晨, 马胜, 王璐, 等. 三维片上网络体系结构研究综述[J]. 计算机学报, 2016, 38(9): 1812-1828.
- [4] 刘有耀. 片上网络拓扑结构与通信方法研究[D]. 西安:西安电子科技大学, 2009.
- [5] Sullivan H, Bashkow T R. A large scale, homogeneous, fully distributed parallel machine[C]// Proceedings of the 4th annual symposium on Computer architecture. Munchen, Germany, ACM Press, 1977: 105-117.
- [6] Valiant L G, Brebner G J. Universal schemes for parallel communication[C]// Proceedings of the thirteenth annual ACM symposium on Theory of computing. New York, NY, USA, ACM Press, 1981: 263-277.
- [7] Nesson T, Johnsson S L. Romm routing on mesh and torus networks[C]// Proceedings of the seventh annual ACM symposium on Parallel algorithms and architectures. Santa Barbara, California, ACM Press, 1995:

275-287.

- [8] Seo D, Ali A, Lim W T, et al. Near-optimal worst-case throughput routing for two-dimensional mesh networks[J]. International Symposium on Computer Architecture, 2005, 33(2): 432-443.
- [9] Ramanujamand R S, Lin B. Randomized partially-minimal routing: near-optimal oblivious routing for 3D mesh networks[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2012, 20(11): 2080-2093.
- [10] Guang Sun, Chia-Wei Chang, Bill Lin. An oblivious routing algorithm for 3D mesh networks to achieve a new worst-case throughput bound[J]. IEEE Embedded Systems Letters, 2012, 4(4): 98-101.

作者简介:

王振江 男, (1990-), 硕士. 研究方向为片上网络.

E-mail: wzjabel@162.com.

段新明 男, (1970-), 博士. 研究方向为互联网络技术、片上网络技术、容错路由.

(上接第 22 页)

慢的不足. 仿真实验结果表明, GWO-BP 算法较传统算法和 GA-BP 算法相比具备收敛速度快, 复原精度高的特点. GWO-BP 复原算法可以满足机载成像装置图像复原需求, 下一步工作拟结合 GWO-BP 神经网络算法与硬件结合实现.

参考文献:

- [1] 殷兴良. 气动光学原理[M]. 北京: 中国宇航出版社, 2003.
- [2] Sutton G, Pond J, Snow R, et al. Hypersonic interceptor performance evaluation center aero-optics performance predictions [J]. Journals of Gerontology, 2013, 66(4): 444-453.
- [3] 张士杰, 李俊山, 杨亚威, 等. 湍流退化红外图像校正算法 [J]. 红外与激光工程, 2014, 43(11): 3670-3675.
- [4] 王锋, 张昆帆, 王希云, 等. TV 正则化的双迭代 RL 湍流退化图像复原算法[J]. 信息工程大学学报, 2014, 15(2): 210-214.
- [5] 朱瑞飞, 魏群, 王超, 等. 基于随机点扩散函数的多帧

湍流退化图像自适应复原方法[J]. 中国光学, 2015, 7(3): 368-377.

- [6] Hecht Nielsen R. Theory of backpropagation neural networks[C]// International Joint Conference on Neural Networks. IEEE Xplore, 1989, 1(1): 593-605.
- [7] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69(3): 46-61.
- [8] 徐辰华, 李成县, 喻昕, 等. 基于 Cat 混沌与高斯变异的改进灰狼优化算法[J]. 计算机工程与应用, 2017, 53(4): 1-9.

作者简介:

杨书杰 男, (1994-), 硕士研究生. 研究方向为气动光学退化效应仿真及退化图像复原. E-mail: jief73@hotmail.com.

叶 霞 女, (1977-), 副教授, 硕士生导师. 研究方向为数据库技术、网络信息安全技术、图像处理.

李俊山 男, (1956-), 教授, 博士生导师. 研究方向为智能图像处理与目标感知识别、网络信息安全、电子对抗模拟与仿真.