

基三分层互连网络和 2-D Mesh 的比较

乔保军^{1,2} 石 峰¹ 计卫星¹

(北京理工大学计算机科学技术学院 北京 100081)¹ (河南大学数据与知识工程研究所 开封 475001)²

摘要 多核处理器(multi-core processor)成为高性能处理器体系结构的研究发展方向,核间的连接方式对多核处理器性能的发挥起着重要作用。从降低节点度、减少网络链路数和缩短网络直径的角度出发,提出了一种用于片上核间互连的新型分层互连网络——基三分层互连网络(THIN),该网络拓扑简单,节点度数低,网络链路数相对较少,并具有明显的层次性和对称性以及良好的扩展性。深入比较了 THIN 和 2-D Mesh 的静态度量和无阻塞延迟,比较结果表明:在网络规模较小时,THIN 比 2-D Mesh 更宜于用来构建片上核间的通信网络。

关键词 多核处理器,片上互连网络,2-D Mesh,网络拓扑

Comparison of the Triplet-based Hierarchical Interconnection Network and 2-D Mesh for Multi-core Processor

QIAO Bao-Jun^{1,2} SHI Feng¹ JI Wei-Xing¹

(School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081)¹

(Institute of Data and Knowledge Engineering, Henan University, Kaifeng 475001)²

Abstract High-performance processor architectures are moving towards designing that feature multiple processing cores on a single chip. Efficient on-chip interconnection network is very essential to the performance of multi-core processor. A new interconnection network for multi-core processor, which is named triple-based hierarchical interconnection network (THIN), is proposed. The topology of the network is very simple and it has obviously hierarchical, symmetric and scalable characteristic. The interconnection network properties and latency in no-blocking mode of THIN and 2-D Mesh are compared. The compare result shows when there are not too many nodes THIN is more fit to construct interconnection network for multi-core processor than 2-D Mesh.

Keywords Multi-core processor, On-chip interconnection network, 2-D Mesh, Network topology

1 引言

随着集成电路制造技术的不断发展,单个硅芯片上集成上亿个晶体管已经成为可能,多核处理器(multi-core processor)成为高性能处理器体系结构的研究发展方向。片上核间的连接方式对多核处理器的性能有着极为重要的影响^[1]。传统的多核处理器一般采用时分多路(TDM, Time-Division Multiplexed)总线作为多核间的通讯方式,如 IBM Core Connect^[2]和 ARM AMBA^[3]。总线方式最大的不足在于它不具有可伸缩性,随着处理器核的增加,共享传输介质将成为通信瓶颈,导致系统性能下降^[4]。

一些研究表明:片上互连网络(on-chip interconnection networks)是一种比总线方式更适于作为核间互连的通讯方式^[5,6]。从图论的角度出发,许多互连网络都具有良好的数学特性,例如:全互连网络、环网(torus)和超立方体(hypercube)等。但是这些互连网络结构比较复杂难以在片上实现,所以目前人们在研究多核处理器的核间互连时,更多的是采用 2-D Mesh 作为片上互连网络^[7]。

本文提出一种用于核间互连的新型片上互连网络——基三分层互连网络(THIN, Triplet-based Hierarchical Interconnection Network),该网络拓扑简单,节点度数低,具有明显的

层次性和对称性以及良好的扩展性。本文详细地比较了 THIN 和 2-D Mesh 网络拓扑的静态度量和无阻塞延迟,比较结果表明:在网络规模不大的情况下,THIN 比 2-D Mesh 更适于用来构建片上核间通信的互连网络。

2 THIN 网络拓扑

THIN 是一种层次化的、可扩展的互连拓扑结构。该结构的第 0 层是单个节点,如图 1(a)所示。通过 3 条通信链路将 3 个节点彼此互连形成一个三角形,从而构成该结构的第 1 层,如图 1(b)所示。1 层网络是构造 THIN 的基本构件,可以将该基本构件看作一个超节点。

利用这个基本构件,通过使用迭代函数族(IFS)可以构造出任意层次的基三分层互连网络。假设该结构的迭代函数族为 $IFS\{F_1, F_2, F_3\}$,将 1 层基三分层互连网络看作是 1 次迭代后的网络 $N(1)$, $N(k)$ 表示 k 次迭代后得到的 k 层网络,那么 THIN 的构造过程就可以表述为:

$$N_{k+1} = \bigcup_{i=1}^3 F_i(N_k) \quad (1)$$

THIN 的递归构造过程为:在基本构件的基础上,将每个节点用一个低层网络替代,从而得到更高层的一个三角形网络结构。重复这一过程,我们可以构造出满足应用需求的任

意层次的 THIN。该网络拓扑结构并不局限于平衡构造,还可以进行非平衡构造。图 1(d)给出了一个经过 3 次迭代,含

有 27 个节点的三层 THIN。

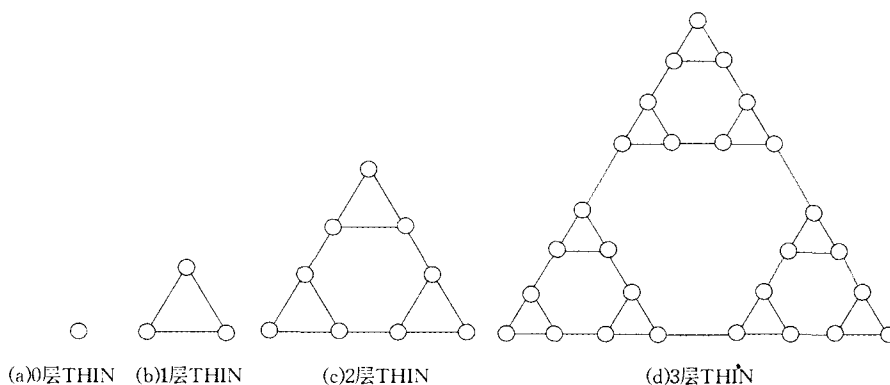


图 1 THIN 互连网络拓扑结构

3 THIN 和 2-D Mesh 的静态度量比较

网络拓扑的静态度量主要包括有:网络的度、链路数和网络直径。

定义 3.1 对于 k 层 THIN,网络的节点集记为 V_k ,网络节点个数为 N ,则

$$N = |V_k| = 3^k \quad (2)$$

定义 3.2 与节点 i 相邻的所有节点的个数称为节点的度 (node degree), 记为 d_i 。图中的最大节点度称之为图的度, 表示为 d :

$$d = \max(d_i) \quad (3)$$

THIN 中网络度表示为 d_{THIN} ,

$$d_{THIN} = 3 \quad (4)$$

定义 3.3 k 层 THIN 的链路总数计为 L_k 。

根据 THIN 网络的迭代构建过程,该网络中的链路总数可以使用如下递推公式表示:

$$\begin{cases} L_1 = 3 \\ L_k = 3L_{k-1} + 3 \end{cases} \quad (5)$$

由递推公式(5)可知:

$$L_k = 3 \times (3^k - 1) / 2 = 3 \times (N - 1) / 2 \quad (6)$$

其中 N 表示 k 层 THIN 中节点的个数。

定义 3.4 连接两个节点 i 和 j 的最短路径中包含的边数称为节点 i 和 j 的距离, 表示为 $P_{i,j}$ 。网络中的任意两节点间的距离的最大值称为网络的直径 P , 即

$$P = \max(P_{i,j}) \quad (7)$$

k 层 THIN 互连网络的网络直径记为 $P_{THIN}(k)$, 根据 THIN 网络的迭代构建过程,该网络直径可以使用如下递推公式表示:

$$\begin{cases} P_{THIN(1)} = 1 \\ P_{THIN(k)} = 2P_{THIN(k-1)} + 1 \end{cases} \quad (8)$$

由递推公式(8)可知:

$$P_{THIN(k)} = 2^k - 1 = 2^{\log_3 N} - 1 \quad (9)$$

表 1 THIN 和 2-D Mesh 的网络拓扑静态度量比较

网络类型	度	链路总数	网络直径
THIN	3	$3 \times (N - 1) / 2$	$2^{\log_3 N} - 1$
2-D Mesh	4	$2(N - \sqrt{N})$	$2(\sqrt{N} - 1)$

表 1 对比了 THIN 和 2-D Mesh 的网络拓扑的静态度

量,其中 N 表示网络节点数。从表 1 中可知,THIN 和 2-D Mesh 的度都是常数,并且 THIN 的度更小。不变的节点度使得网络接口的开销不会随着网络规模的变化而发生改变,适合 VLSI 实现,网络更容易扩展。而且由于较小的节点度,在通道受到硬件布线限制时可以有较大的线宽,对网络的带宽有利。

数据链路数表示了构造网络的成本和网络的复杂度,当节点数增加时,为了使连接代价达到最小,总链路数应当按照线性规律增加。在相同节点下,THIN 的链路总数也比 2-D Mesh 的少,这一点对于构建核间互连网络而言非常重要,因为较少的链路数占用的片上资源也相对较少。

网络直径是互连网络的一重要参数,它直接影响到节点间的通信延迟,通常在包交换网络中要求网络具有尽可能小的直径。图 2 比较了 THIN 和 2-D Mesh 在不同节点数下的网络直径的取值情况。从图中可知,当网络规模不大时,THIN 的网络直径比 2-D Mesh 的要小。

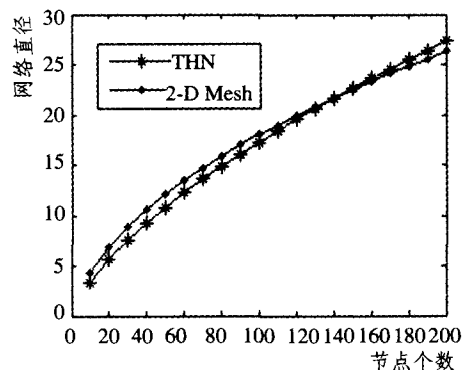


图 2 THIN 和 2-D Mesh 的网络直径值比较图

4 THIN 和 2-D Mesh 的无阻塞延迟比较

网络拓扑的静态度量主要反映网络的拓扑性质,为了对网络性能进行全面的评价,还必须考虑网络通信延迟这一重要的动态度量,本小节主要比较了 THIN 和 2-D Mesh 网络的无阻塞延迟。

假设消息包含 L 位数据,物理微片和微片大小相等,都等于物理数据通道的宽度 W 位,消息头假设为一个微片,这样消息的大小为 $L + W$ 位^[8]。文[8]给出了互连网络在虚切入交换机制(virtual cut-through)下的无阻塞通信延迟的计算

公式:

$$T = D_{avg}(t_r + t_s + t_w) + \max(t_s, t_w) \lceil \frac{L}{W} \rceil \quad (10)$$

其中, t_r 表示作出路由决策所花费的时间, t_s 为路由器内部延迟或交换延迟, 即 W 位的微片从路由器的输入传送到输出的时间, t_w 是通道延迟, D_{avg} 是互连网络的平均距离。

定义 4.1 互连网络的平均距离定义为网络中任意两点之间的最短路径长度之和与路径总数之比, 这里包括节点到自身的长度为 0 的路径^[9], 记为 D_{avg} , 其计算公式如下:

$$D_{avg} = \sum_{i=0}^P i \times \rho(i) \quad (11)$$

其中 P 表示网络的直径, $\rho(i)$ 表示距离为 i 的消息占消息总量的比例。

根据 THIN 互连网络的拓扑结构, k 层 THIN 互连网络的平均距离 D_{THIN} 的计算公式为 (详细的推导过程限于篇幅从略):

$$D_{THIN} = \frac{1}{3^{k-1}} + \frac{16 \times (6^{k-1} - 1)}{5 \times 3^k} - \frac{1}{3} \quad (12)$$

文[10]给出了求解 2D-Mesh 的平均距离 $D_{2-D Mesh}$ 的计算公式:

$$D_{2-D Mesh} = \frac{2(N-1)}{3\sqrt{N}} \quad (13)$$

在比较具有不同度的互连网络的平均距离时, 由于网络度越高则越有利于降低平均距离, 因而只有平均距离的定义是难以对通信延迟做出科学的分析, 因此引入归一化平均距离的概念^[11]。

定义 4.2 互连网络的归一化平均距离 μ 定义为网络的平均距离 (D_{avg}) 和网络度 (d) 的乘积:

$$\mu = d D_{avg} \quad (14)$$

由公式(12)可得 k 层 THIN 互连网络的归一化平均距离 μ_{THIN} 为:

$$\mu_{THIN} = 3 \times \left(\frac{1}{3^{k-1}} + \frac{16 \times (6^{k-1} - 1)}{5 \times 3^k} - \frac{1}{3} \right) \quad (15)$$

由公式(13)可知二维网络的归一化平均距离 $\mu_{2-D Mesh}$ 为:

$$\mu_{2-D Mesh} = \frac{8(N-1)}{3\sqrt{N}} \quad (16)$$

在比较互连网络的无阻塞延迟时, 用网络的归一化平均距离 (μ) 替代公式中的平均距离 (D_{avg})。由公式(10)和(15)可以得出 k 层 THIN 互连网络的无阻塞延迟计算公式为:

$$\begin{aligned} T_{THIN} &= 3 \times \left(\frac{1}{3^{k-1}} + \frac{16 \times (6^{k-1} - 1)}{5 \times 3^k} - \frac{1}{3} \right) \times (t_r + t_s + t_w) \\ &\quad + \max(t_s, t_w) \lceil \frac{L}{W} \rceil \\ &= 3 \times \left(\frac{3}{N} + \frac{16 \times (6^{\log_3 N - 1} - 1)}{5 \times N} - \frac{1}{3} \right) \times (t_r + t_s + \\ &\quad t_w) + \max(t_s, t_w) \lceil \frac{L}{W} \rceil \end{aligned}$$

其中, N 是 k 层 THIN 互连网络中节点的个数。

由公式(10)和(16)可以得出含有 N 个节点的二维网格的无阻塞延迟计算公式为:

$$T_{2-D Mesh} = \frac{8(N-1)}{3\sqrt{N}} (t_{hr} + t_s + t_w) + \max(t_s, t_w) \lceil \frac{L}{W} \rceil \quad (18)$$

图 3 比较了在采用相同的路由决策机制、相同的消息交

换机制和相同的通信带宽的条件下, THIN 和 2-D Mesh 的无阻塞延迟情况。在这里假设路由决策时间 (t_r)、交换延迟 (t_s) 和通道延迟 (t_w) 都是一个固定常量, 我们用不同节点个数下的网络归一化平均距离来度量网络的无阻塞延迟。从图中可知, 当网络规模 (节点个数) 不是很大时, THIN 的无阻塞延迟要比 2-D Mesh 的小。

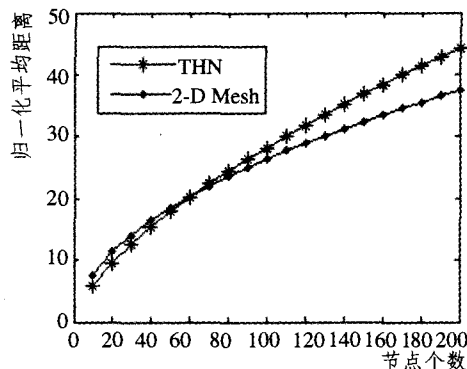


图 3 THIN 和 2-D Mesh 的无阻塞延迟比较图

结束语 本文提出了一种新型的片上互连网络——基三分层互连网络 (THIN), 该网络拓扑简单、节点度数低, 具有明显的层次性和对称性以及良好的扩展性。另外, THIN 中的最底层节点之间采用全互连结构, 较高层次网络之间的连接链路相对较少, 降低了网络复杂度, 减少网络本身对片上资源的占用。在综合考虑网络度、数据链路、网络直径和无阻塞延迟的情况下, 当处理器核较少时, THIN 比 2-D Mesh 更宜于用来构建核间的通信网络。

参考文献

- 1 Parcerisa J M, Sahuquillo J. On-chip interconnects and instruction steering schemes for clustered microarchitectures. *IEEE Trans. on Parallel and Distributed Systems*, 2005, 16(2): 130~144
- 2 IBM CoreConnect Bus Architecture. <http://www-03.ibm.com/chips/products/coreconnect/index.html>, 2005-04-15
- 3 ARM AMBA. <http://www.arm.com/products/solutions/AMBAHomePage.html>, 2005-04-13
- 4 Wiklund D, Liu Dake. Design of a system-on-chip switched network and its design support. In: *IEEE 2002 Int. Conf. on Communications, Circuits and Systems and West Sino Expositions*, IEEE Press, 2002, 2: 1279~1283
- 5 Dally W J, Towles B. Route packets, not wires: on-chip interconnection networks. In: *Proc. of the 38th Design Automation Conference*, Las Vegas, NV, June 2001. 681~689
- 6 Benini L, Micheli G D. Networks on chips: a new SOC paradigm. *IEEE Computer*, 2002, 35(1): 70~78
- 7 Kumar S, et al. A network on chip architecture and design methodology. In: *IEEE Computer Society Annual Symposium on VLSI*, April 2002. 117~124
- 8 Duato J, Yalamanchili S, Ni L. 并行计算机互连网络技术: 一种工程方法. 谢伦国, 张民选, 窦强, 等译. 北京: 电子工业出版社, 2004
- 9 Dandamudi S P. Hierarchical Interconnection Networks for Multicomputer Systems. *IEEE Transactions on Computers*, 1990, 39(6): 786~797
- 10 董迎飞, 王鼎兴, 郑纬民. 精确计算 n 维 Mesh 网络和 n 维 Torus 网络的平均最短路径长度. *计算机学报*, 1997, 20(4): 376~380
- 11 高鹏. 并行数字信号处理机中的结点互连技术研究: [博士论文]. 哈尔滨: 哈尔滨工程大学, 2004