

基于双层片上网络的路由算法

冯 杰, 荆明娥, 虞志益

(复旦大学专用集成电路与系统国家重点实验室, 上海 201203)

摘 要: 针对新型双层片上网络结构的路由问题, 提出一种利用片上网络资源的快速路由算法。将核间通信依据通信频率和总量大小划分成 4 类, 根据链路交换网络适用于大数据量核间通信的特点, 优先将通信总量大且通信频率低的核间通信固化在链路交换网络中, 再利用快速布线算法在最少占用片上网络资源的情况下解决路径配置问题。实验结果表明, 与模拟退火算法相比, 该算法的运行时间减少 74%, 链路交换网络资源利用率提高 21%, 应用程序性能提升 16% 以上。

关键词: 多核处理器; 双层片上网络; 链路交换网络; 包交换网络; 路由算法

中文引用格式: 冯 杰, 荆明娥, 虞志益. 基于双层片上网络的路由算法[J]. 计算机工程, 2016, 42(4): 83-87.

英文引用格式: Feng Jie, Jing Ming'e, Yu Zhiyi. Routing Algorithm Based on Double-layer Network on Chip [J]. Computer Engineering, 2016, 42(4): 83-87.

Routing Algorithm Based on Double-layer Network on Chip

FENG Jie, JING Ming'e, YU Zhiyi

(State Key Laboratory of ASIC and System, Fudan University, Shanghai 201203, China)

[Abstract] This paper presents a fast routing algorithm for the new double-layer Network on Chip (NoC) to make full use of the NoC resources. According to the frequency and total volume of the inter-core communications, the proposed algorithm firstly divides them into four types. Considering that circuit network is much more suitable for communications of large volume, the proposed algorithm assigns these communications to circuit network layer, and then uses the fast routing method to configure the paths for remaining communications in packet network layer. Experimental result shows that the proposed algorithm uses 74% less running time and results in 21% more resource usage of circuit network than simulated annealing algorithm. In addition, this algorithm can improve application performance by more than 16%.

[Key words] multi-core processor; double-layer Network on Chip (NoC); link exchange network; packet exchange network; routing algorithm

DOI: 10.3969/j.issn.1000-3428.2016.04.016

1 概述

半导体工艺的不断发展使晶体管集成度逐步增加, 为多核处理器提供了制造基础。在单核处理器性能提高遭遇瓶颈的背景下, 多核处理器已然成为主流^[1], 并且片上集成的处理器核数不断增加, 如文献[2-4]所分别实现的 16 核、48 核和 64 核处理器。在这样的趋势下, 片上网络 (Network on Chip, NoC) 渐渐取代基于总线的片上系统 (System on Chip, SoC) 成为众核处理器的主流互联方式^[5-6]。然而, 这也为基于此类硬件的并行程序开发带来了挑战, 如何充分调用众多处理器核成为重

要课题。

学术界通常将并行过程分为任务划分 (Partitioning)、调度 (Scheduling)、映射 (Mapping) 和路由 (Routing) 4 个部分^[7]。其中, 路由确定了不同处理器核间通信的方式和路径, 对片上网络通信的效率、吞吐率和能耗有至关重要的影响。如何根据硬件所提供的资源确定路由成为一个复杂的问题: (1) 处理器核数众多; (2) 片上网络结构众多, 负责片上网络通信的路由器 (Router) 存在基于链路交换和基于包交换等多种方式^[8-9]; (3) 需要考虑除片上网络通信的其他通信方式, 如共享存储器; (4) 当片上网络中某些节点出现故障时, 需要考虑相应的容

基金项目: 国家自然科学基金资助项目 (61103008, 61202263)。

作者简介: 冯 杰 (1988 -), 男, 硕士研究生, 主研方向为多核片上网络路由算法; 荆明娥, 副研究员; 虞志益, 副教授。

收稿日期: 2015-03-05 **修回日期:** 2015-04-29 **E-mail:** 11212020013@fudan.edu.cn

错问题。

为解决并行编程中的路由问题,本文针对文献[10]所实现的24核双层片上网络处理器,提出一种基于快速布线算法的解决方法,并对该种算法的效率和实际应用进行分析。

2 双层片上网络

文献[10]的片上网络结合了传统的包交换网络和链路交换网络,采用双层二维网格结构,既有包交换网络也有链路交换网络,并用包交换网络来配置链路交换网络,如图1所示^[8]。

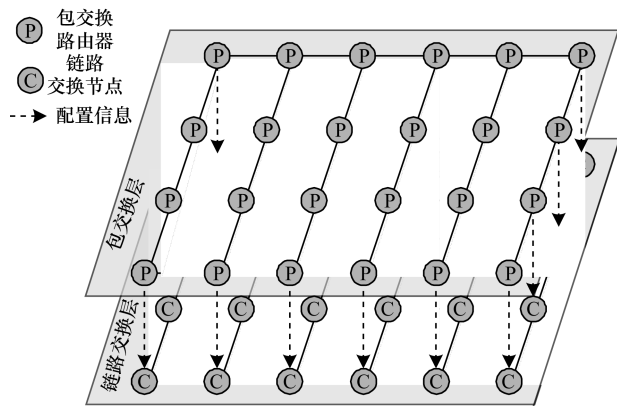


图1 双层片上网络结构

包交换网络具有优异的可扩展性、良好的灵活性、高效的资源利用率等优点。链路交换网络灵活性相对较差,但对比包交换网络具有更高的能量效率,并且当核间通信量比较大或者传输路径比较长时,链路交换网络在传输延迟上有明显的优势。

在硬件自身具备的路由功能方面,包交换网络提供固定的X-Y路由算法。链路交换网络节点路由器(即链路交换路由器)的结构类似于多路选择器,其5个端口(分别是东西南北4个连接其他链路交

换路由器的端口以及链路交换路由器与对应处理器核连接的端口)之间的连接与释放通过与之连接的包交换路由器进行配置,同一个端口在同一时间只能与一个其他端口进行连接。在使用链路网络时,还需要注意不能配置出环路,否则,电路就会发生振荡,处于不稳定状态。图2展示了链路交换路由器可行的配置方式(图2中虚线)以及不可行的配置方式(图2中点虚线)。

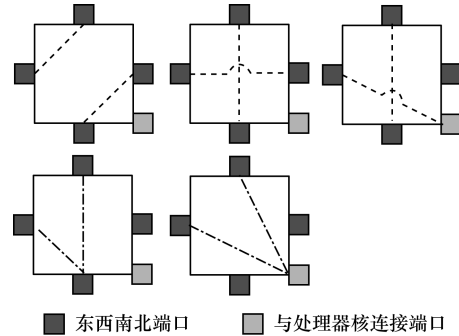


图2 链路路由器配置示意图

3 基于快速布线算法的路由算法

3.1 核间通信图

通过软件工具对串行程序进行划分、调度和映射处理,或者通过手动划分映射。在多媒体和通信等应用程序被“切割”后分别映射到多核处理器不同的处理器核上。图3给出了将长期演进(Long Term Evolution, LTE)程序手动划分映射到24个处理器核上的结果示意图。在应用程序被映射完成后,不同处理器核之间的通信关系随之确定,核间通信图就是用来描述不同处理器核之间通信方向、通信次序与通信量大小的矢量图,它包含了核间通信的所有信息。图4即是图3中LTE程序的核间通信图。

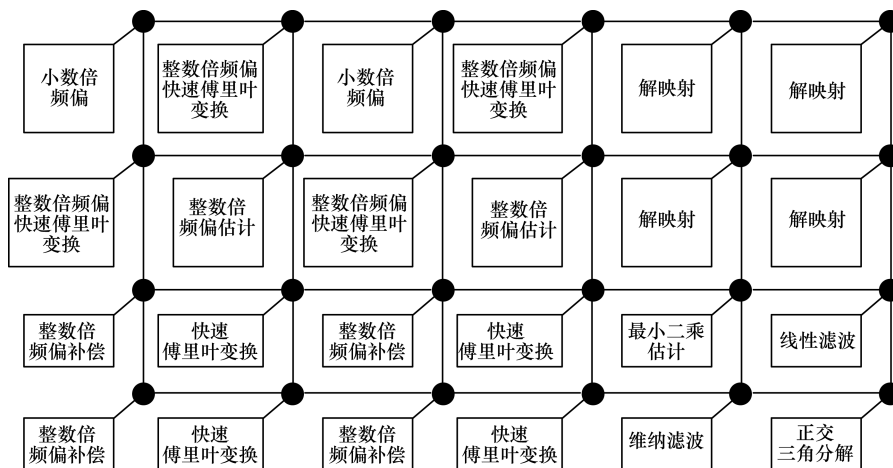


图3 LTE程序手动划分映射到24核处理器的结果示意图

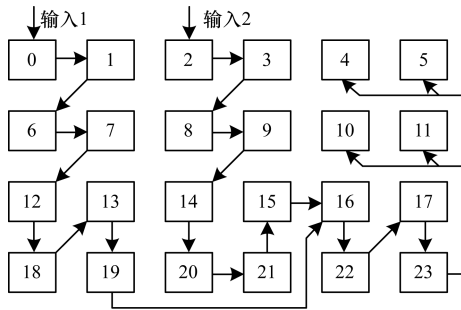


图 4 核间通信图示意图

3.2 核间通信分类

路由解决的是为核间通信找到通信路径并确定具体的传输方式。由于存在双层片上网络,每一层网络都可以用于核间通信,因此需要对核间通信进行分类,以确定每一个通信是经过包交换网络还是经过链路交换网络。根据 2 个不同处理器核之间通信的频率与通信总量的大小 2 个维度,可以将核间通信分为 4 类:(1)总通信量大且通信频率低;(2)总通信量大且通信频率高;(3)总通信量小且通信频率低;(4)总通信量小且通信频率高。

根据对包交换网络和链路交换网络特性的分析可知,总通信量大的核间通信适合通过链路交换网络进行传输,而总通信量小的核间通信则适合通过包交换网络进行传输,并且,当通信总量差不多的时候,通信频率低的核间通信由于单次平均通信量大,更适合通过链路交换网络进行传输。这样,既可以发挥包交换网络良好的灵活性和高效的资源利用性,又可以发挥链路交换网络能量效率高的特性。最后,考虑到通信及多媒体应用核间通信关系相对简单而通信量大的特点,可以将链路交换网络进行通信的核间通信路径固化下来,即每条核间通信路径只建立一次而永不释放。由于链路交换网络核间通信路径的建立与释放有一定的时间代价和能量代价,固化路径可以减少通信路径建立与释放的次数,从而提高核间通信的效率同时降低功耗。

3.3 算法描述

3.3.1 链路交换网络路由处理

就单层链路交换网络而言,由于链路交换路由器的端口在同一时间只能与一个其他端口连接,当在链路交换网络层的通信路径只建立一次而不动态地建立、释放的时候,在链路交换网络层为核间通信进行路由配置的问题就转变成了为核间通信寻找互不冲突的通信路径的问题。针对这一问题,本文参考文献[11-12]针对现场可编程逻辑门阵列(Field Programmable Gate Array, FPGA)的快速布线算法设

计了一种对冲突敏感的快速为核间通信匹配通信路径的算法如下:

```

While congestion exists/* global routing */
  Rip up all routing trees
  For each communication pair (si, ti)/* single communication routing */
    Initialize priority queue PQ with si at cost of 0
    While the target ti has not been reached
      Remove lowest cost node m from PQ
      Loop over all neighbors k of node m
        Add k to PQ at cost of ck + Pim
      End Loop
    End While
    Loop over all nodes k in path ti to si/* backtrace */
      Update ck
      Add k to routing tree RTi
    End Loop
  End For
End While
    
```

其中,某一路径 i (起点 s_i , 终点 t_i) 占用链路交换路由器 k 的代价函数 c_k 由下式决定, P_{im} 是从起点 s_i 到链路交换路由器 m 的总代价。

$$c_k = (b_k + h_k) \times p_k + \Delta D$$

其中, b_k 是占用链路交换路由器 k 的基础代价; h_k 是当前全局路由循环之前的所有全局路由循环中占用链路交换路由器 k 的历史冲突的函数; p_k 是当前全局路由循环中除路径 i 外占用链路交换路由器 k 的通信路径数目的函数; ΔD 是链路交换路由器 k 与目标链路交换路由器 t_i 的曼哈顿距离对比链路交换路由器 k 的前继链路交换路由器 m 与目标链路交换路由器 t_i 的曼哈顿距离的变化的函数,若通信路径 i 在搜索路径的过程中占用链路交换路由器 k 之后离目标链路交换路由器 t_i 的距离变得更近,则 ΔD 的值为负,反之为正。本文在算法设计的时候,特意将离中心远的链路交换路由器的 b_k 值设置的更小,这样在选择路径的时候会优先考虑边缘位置的链路交换路由器,从而减少更容易发生冲突的处于中间位置的链路交换路由器出现冲突的概率,提高算法的时间效率以及路径匹配的成功率。

3.3.2 双层片上网络路由处理

根据上文的介绍,当对双层片上网络路由进行全局考虑时,首先应读入核间通信图,获取核间通信的全部信息,然后对核间通信进行分类,接着根据分类结果确定在链路交换网络进行通信的核间通信,并为之配置合理的通信路径,最后将其余的核间通信安排在包交换网络,具体算法流程如图 5

所示。

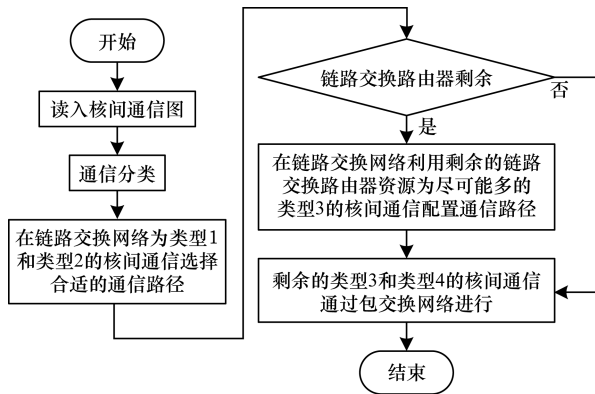


图5 基于双层片上网络的快速路由算法

4 性能评估

4.1 与模拟退火算法的比较

模拟退火算法因其求解的高效性、高质量性和容易实现的特点,被广泛应用于组合优化问题的求解。传统路由算法研究的是路由器路由策略的设计,而本文路由算法研究的则是合理利用路由器资源以实现高效的核间通信。再加上本文针对的是一种新型的双层片上网络结构,因而并没有相似研究,在此情况下,本文选择模拟退火算法作为比较算法,比较结果如表1和表2所示。其中,运行时间是2种算法分别在20次随机产生的核间通信图作为测试输入下的平均运行时间,链路交换通信数是被安排在链路交换网络的核间通信数目的平均值,该值越大,表明链路交换网络资源被利用的越充分。

表1 24核处理器下2种算法比较

算法	核间通信数	运行时间/min	链路交换通信数
本文算法	20	1.2	13.6
模拟退火算法	20	4.6	10.8

表2 48核处理器下2种算法比较

算法	核间通信数	运行时间/min	链路交换通信数
本文算法	40	3.9	22.4
模拟退火算法	40	27.3	15.3

从表1和表2的数据对比中可以发现,本文算法对比基于模拟退火算法的路由算法有着明显的优势,在24核下运行时间减少74%,链路交换网络资源利用率提高21%;在48核下运行时间减少86%,链路交换网络资源利用率提高32%。需要注意的是,此处的48核处理器是为评估算法性能而虚拟出来的,并没有真正实现。

4.2 实例分析

通过文献[7]所实现的软件工具链将12点快速傅里叶变换(Fast Fourier Transformation, FFT)串行程序划分、调度、映射到文献[10]所实现的24核处理器上,再通过本文算法实现对映射好的并行程序的路由配置,便可将处理过的并行程序在24核处理器上进行仿真。如图6所示,采用本文算法相对于只用包交换网络和只用链路交换网络进行通信,12点FFT程序整体性能分别提高16%和34%。

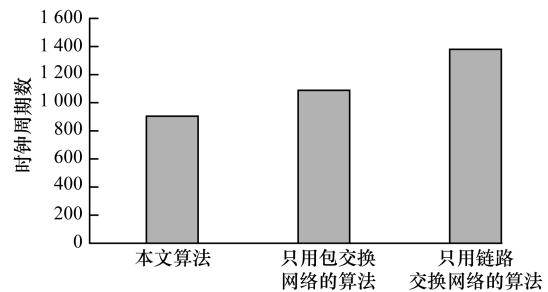


图6 不同路由算法下12点FFT时钟周期数

类似地,将LTE同步模块手动映射到12个处理器核上,如图7所示,再利用本文算法对核间通信进行配置,最后进行仿真。图8的结果表明,采用本文算法相对于只用包交换网络和只用链路交换网络进行通信,LTE同步模块程序整体性能分别提高23%和17%。

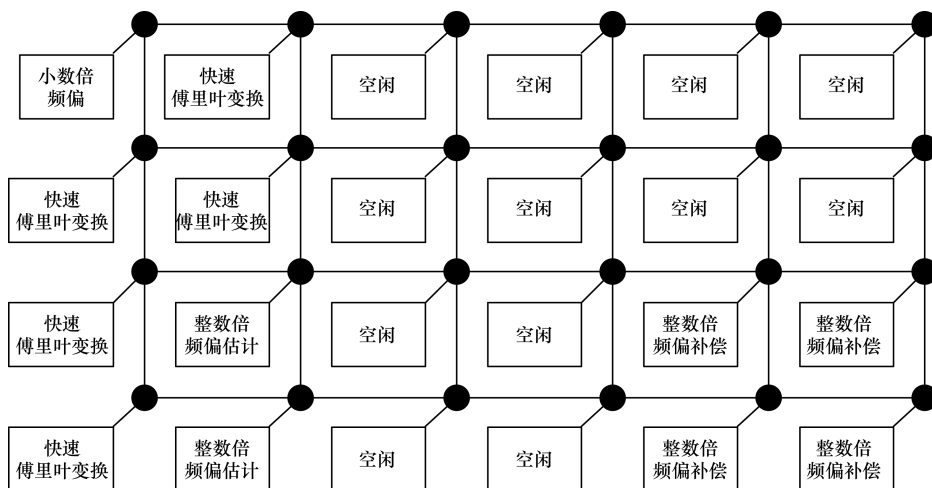


图7 LTE同步模块映射结果示意图

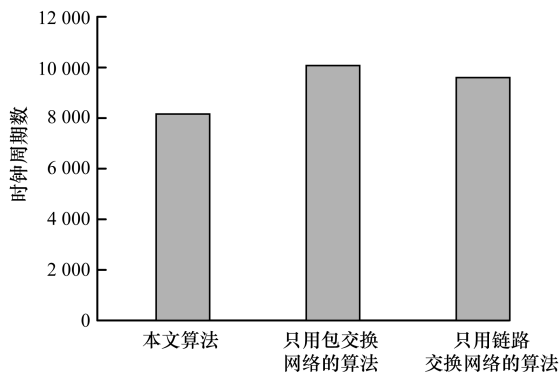


图 8 不同路由算法下 LTE 同步模块的时钟周期数

5 结束语

为了合理利用双层片上网络,提高系统吞吐率和应用程序性能,本文提出一种快速路由算法。通过对核间通信进行分类,为核间通信确定合理的通信方式,并采用基于 FPGA 布线算法的方法为使用链路交换网络进行通信的核间通信选择合适的通信路径。实验结果表明,该算法能够有效提高通信效率与应用程序性能。下一步工作将研究共享存储器在核间通信中的作用,考虑共享存储器与片上网络的结合。

参考文献

- [1] Majed C, Sami Y, Claude T, et al. A Hypercube-based NoC Routing Algorithm for Efficient All-to-all Communications in Embedded Image and Signal Processing Applications [C]//Proceedings of International Conference on High Performance Computing and Simulation. Washington D. C. , USA: IEEE Press, 2010:623-630.
- [2] Yu Zhiyi, You Kaidi, Xiao Ruijin, et al. An 800 MHz 320 mW 16-core Processor with Message-passing and Shared-memory Inter-core Communication Mechanisms[C]//Proceedings of IEEE International Solid-state Circuits Conference. Washington D. C. , USA: IEEE Press, 2012:64-66.
- [3] Mattson T G, Wijngaart R F V D, Riepen M, et al. The 48-core SCC Processor: The Programmer's View[C]//Proceedings of ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis. New York, USA: ACM Press, 2010:1-11.
- [4] Bell S, Edwards B, Amann J, et al. TILE64 Processor: A 64-core SoC with Mesh Interconnect[C]//Proceedings of IEEE International Solid-state Circuits Conference. Washington D. C. , USA: IEEE Press, 2008:88-598.
- [5] 俞剑明. 容偏差百核片上网络设计与多核粒度建模[D]. 上海: 复旦大学, 2014.
- [6] Hu Jingcao, Marculescu R. Energy and Performance-aware Mapping for Regular NoC Architectures[J]. IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems, 2005, 24(4): 551-562.
- [7] 秦臻南. 程序并行划分算法及自动并行工具框架实现[D]. 上海: 复旦大学, 2013.
- [8] 欧 鹏. 多核处理器关键技术研究——单核及核间通讯的架构与实现[D]. 上海: 复旦大学, 2013.
- [9] Chang Kuei-Chung, Shen Jih-Sheng, Chen Tien-Fu. Evaluation and Design Trade-offs Between Circuit-switched and Packet-switched NoCs for Application-Specific SoCs[C]//Proceedings of Design Automation Conference. New York, USA: ACM Press, 2006: 143-148.
- [10] Ou Peng, Zhang Jiajie, Quan Heng, et al. A 65nm 39GOPS/W 24-Core Processor with 11Tb/s/W Packet Controlled Circuit-switched Double-layer Network-on-Chip and Heterogeneous Execution Array [C]//Proceedings of IEEE International Solid-state Circuits Conference. Washington D. C. , USA: IEEE Press, 2013: 56-57.
- [11] Swartz J, Betz V, Rose J. A Fast Routability-driven Router for FPGAs [C]//Proceedings of the 6th International Symposium on FPGAs. New York, USA: ACM Press, 1998:140-151.
- [12] Ebeling C, McMurchie L, Hauck S A, et al. Placement and Routing Tools for the Triptych FPGA [J]. IEEE Transactions on Very Large Scale Integration Systems, 1995, 3(4): 473-482.
- [7] Salsano S, Detti A, Cancellieri M, et al. Transport-layer Issues in Information Centric Networks[C]//Proceedings of the 2nd ICN Workshop on Information-centric Networking. New York, USA: ACM Press, 2012:19-24.
- [8] Rozhnova N, Fdida S. An Effective Hop-by-hop Interest Shaping Mechanism for CCN Communications [C]//Proceedings of IEEE Conference on Computer Communications Workshops. Washington D. C. , USA: IEEE Press, 2012: 322-327.
- [9] Carofiglio G, Gallo M, Muscariello L. Joint Hop-by-hop and Receiver-driven Interest Control Protocol for Content-centric Networks [C]//Proceedings of the 2nd ICN Workshop on Information-centric Networking. New York, USA: ACM Press, 2012:37-42.
- [10] Saino L, Cocora C, Pavlou G. CCTCP: A Scalable Receiver-driven Congestion Control Protocol for Content Centric Networking [C]//Proceedings of IEEE International Conference on Communications. Washington D. C., USA: IEEE Press, 2013:3775-3780.
- [11] 唐 潇, 任勇毛, 李 俊, 等. 一种基于显式反馈的内容中心网络 NDN 拥塞控制算法[J]. 科研信息化技术与应用, 2014, 5(3): 68-77.
- [12] 3GPP. Quality of Service (QoS) Concept and Architecture[EB/OL]. [2015-03-03]. <http://www.3gpp.org/Dyna-Report/23107.htm/>.

编辑 刘 冰

(上接第 82 页)

编辑 刘 冰