

面向应用的 NoC 带宽感知路由技术

林红君,王长山

(西安电子科技大学 计算机学院,西安 710071)

(lemonlhj@163.com)

摘要:片上互连网络是片上通信问题的有效解决方案,但存在严重的资源限制。标准拓扑结构难以满足应用的流量需求,同时还导致大量功耗和面积的开销。适用于通用系统的 NoC 设计难以满足面向服务质量可预测的互连。给出一种面向应用的带宽感知路由技术,针对具体的应用,首先使用基于遗传算法的映射技术获得 IP 核到网络节点的最佳映射,然后通过带宽感知的路由算法为网络中的每条数据传输生成最短路由,并通过虚信道静态分配保证该路由是无死锁的。为了减少路由表的硬件开销,还结合使用了路由表压缩的方法。仿真结果表明,所提出的路由技术与现有的路由算法相比,具有更好的时延性能。

关键词:片上网络;路由;映射;虚信道;流量特征

中图分类号: TP302 **文献标志码:** A

Application-oriented NoC bandwidth aware routing technology

LIN Hong-jun, WANG Chang-shan

(School of Computer Science and Technology, Xidian University, Xi'an Shaanxi 710071, China)

Abstract: The Networks-on-Chip (NoC) approach was proposed as a promising solution to on-chip communication problems, but it is by far more resource limited. The standard topology structure cannot satisfy the application traffic demand and leads to overconsumption of the power and area. These designs that are adequate for general-purpose systems cannot satisfy quality-of-service oriented, predictable interconnects. This paper proposed a new Bandwidth-Aware Routing Technique (BART) that optimized the network performance for application-specific NoCs. More precisely, the specific application and traffic characteristics were given. The proposed routing technique has two phases. In the first phase, a communication-aware mapping technique was used to obtain a near-optimal assignment of IPs to network nodes. In the second phase, a bandwidth-aware routing algorithm was used to find the minimal route for each flow in the network. It was ensured that the routes were deadlock-free through static virtual channel assignment. Besides, BART combined a method for router table compression to reduce the hardware overhead. The evaluation results show that BART has better performance than the existing routing algorithms.

Key words: Network-on-Chip (NoC); routing; mapping; virtual channel; traffic characteristic

0 引言

随着芯片上集成的 IP 数量的增加,基于总线的设计已经不能满足 IP 间的通信要求。片上互连网络 (Network-on-Chip, NoC) 成为片上通信问题的有效解决方案^[1-4]。

与标准的数据宏观网络相比,片上互连网络的设计面临不同的挑战。首先,片上互连网络存在更严重的资源限制。因此,尽可能地降低其面积和功耗开销是一个重要的设计目标;其次,对大多数无线的以及多媒体的片上系统 (System on Chip, SoC) 平台 (如 TI OMAP^[5]) 而言,在设计系统时就可以将任务静态 (或半静态) 地映射到处理器与 IP 核^[6]。因此,在设计时可以分析大多数 SoC 的通信模式和流量特征并为特定的应用定制片上网络。

路由技术是影响片上网络的面积、时延、吞吐和功耗等参数的一个重要因素。文献[7]提出了一种面向应用的路由算法 (Application-Specific Routing Algorithm, APSRA), 利用通信的网络节点对的信息最大化路由函数的适应度并保证是无死

锁的。然而,APSRA 没有考虑不同的通信任务间的带宽需求,导致数据负载在网络中分布不均。

本文提出的带宽感知路由技术 (Bandwidth-Aware Routing Technique, BART) 是针对特定应用的通信模式和流量特征设计的,结合了基于遗传算法的映射技术与带宽感知的路由算法。首先根据对应用的带宽需求的粗略估计,通过基于遗传算法的映射技术获得 IP 核到网络节点的近最优映射。然后通过带宽感知的路由算法为网络中的每条数据传输生成最短路由,达到最小化最大信道负载的目的;同时使用虚信道静态分配保证该路由是无死锁的。最后,为了减少路由表的面积与硬件开销,还结合了路由表压缩的方法,将路由表按照颜色分成不同的群。该路由技术充分利用了网络带宽,使可用的网络带宽满足应用的流量需求,可以应用于规则或不规则的拓扑结构中。

1 基于遗传算法的映射技术

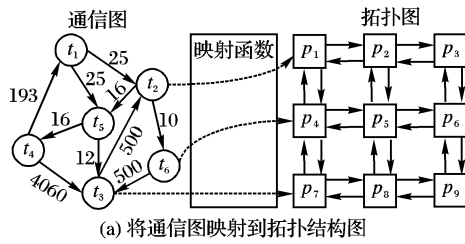
有向图 $G(T, C)$ 构成了一张通信图,其中的每一个节点

收稿日期:2010-06-28;修回日期:2010-08-06。

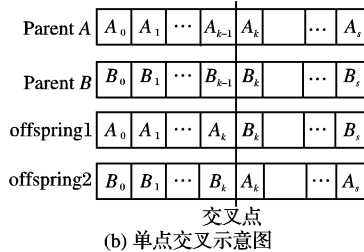
作者简介:林红君(1982-),女,四川广安人,硕士研究生,主要研究方向:面向应用的片上网络;王长山(1950-),男,陕西西安人,副教授,主要研究方向:计算机网络。

$t_i \in T$ 代表每个通信任务,有向边 (t_i, t_j) 表示从 t_i 到 t_j 的数据通信。有向图 $G(P, L)$ 构成 NoC 拓扑结构图,其中的每一节点 $p_i \in P$ 代表网络中的一个节点,有向边 (p_i, p_j) 表示从 p_i 到 p_j 的通信链路。

从通信图 $G(T, C)$ 到拓扑结构图 $G(P, L)$ 的映射关系如图 1(a) 所示。从 $G(T, C)$ 到拓扑结构图 $G(P, L)$ 的映射是一一对应的,映射函数 $M: T \rightarrow P, \text{map}(t_i) = p_j, \forall t_i \in T, \exists p_j \in P$ 。



(a) 将通信图映射到拓扑结构图



(b) 单点交叉示意图

图 1 拓扑结构和单点交叉示意图

文献[8]中提出用映射系数来评估每种映射的优劣程度。映射系数 $Mc = \sum_{i=1}^N \sum_{j=1}^N com_{ij} \times d_{m_i m_j}$, 其中: com_{ij} 表示映射

到节点 i 与节点 j 的任务间传输的字节数,即节点 i 与 j 之间的带宽($weight_{ij}$); m_i 与 m_j 分别表示 IP 核 i 与 j 被映射到网络节点 m_i 与 m_j , $d_{m_i m_j}$ 则表示节点 m_i 与 m_j 之间的等效距离,其计算方法与计算串/并联电路的电阻值相同。 Mc 评估的是当前映射下,应用的所有任务交换的总字节数之和,即总开销。为方便计算,将映射系数修改为: $Mc = \sum_{i=1}^N \sum_{j=1}^N weight_{ij} \times d_{m_i m_j}$ 。

关键的问题是要找到一种映射方式,使 $G(T, C)$ 映射到

$G(P, L)$, 满足 $\min \left\{ \sum_{i=1}^N \sum_{j=1}^N weight_{ij} \times d_{m_i m_j} \right\}$ 。文献[8]中采用随机查找的方法寻找使 Mc 值最小的最优解。然而,随着芯片上集成的 IP 核日趋增多,随机查找的方法将消耗大量执行时间,同时它所求得的最优解仅限于局部最优。

本文对随机查找的方法进行改进,基于遗传算法寻找最优解。算法的实现包括四种操作:选择、交叉、变异及替换。首先,用随机生成的映射方案作为初始种群,初始种群含 S 个个体成员(或称为染色体)。每一条染色体代表一种映射方案。每条染色体由一串基因组成,每个基因代表拓扑结构图中的一个节点,它的值即是映射到该节点的通信图中的节点编号。

采用单点交叉的方法生成后代,随机选择一个交叉点,两个后代交叉点之前的基因分别继承双亲的交叉点之前的基因,交叉点之后的基因以互补的方式继承。交叉方法如图 1(b) 所示。变异即是一条染色体中单个基因之间互换,一条染色体中基因变异的概率为 0.5。计算每个个体成员的适应度函数:

$$f = \sum_{i=1}^N \sum_{j=1}^N weight_{ij} \times d_{m_i m_j}$$

将种群中适应度函数值最小的个体排除,用新生成的后代替换。

算法的终止条件为按上述方法循环执行设定的次数(例如,生成 100 代或 1000 代),直到适应度函数的值不再减小。遗传算法能够在较短的时间进行全局寻优,可以选择最后一代中适应度函数较好的个体作为近最优解。

2 带宽感知的路由算法

2.1 路由器结构

本文所提出的路由技术很大程度上独立于网络拓扑结构和流控机制,因此可以使用典型的虚信道路由器,其结构图见文献[13]。

在面向应用的 NoC 中,为了最大化吞吐量,路由算法将利用从源节点到目的节点的任意路径。因此,路由器必须是可编程的,使得对于每个流,路由可以依据应用配置,并灵活支持任意的路由路径[9]。

因此,面向应用的 NoC 的路由器适合于采用基于表实现的路由器。基于表实现的路由器因其具有可写的内存,允许重配置(或动态重配置)路由算法来高效地处理应用在运行中通信要求的改变。本文采用基于表的源路由。

2.2 问题描述

首先给出流图的定义。

定义 1 设有一张流图 $G(V, E)^{[9]}$, 其中有向边(或称为链路) $(u, v) \in E$, d 为该边上的带宽。为 G 中的每条边赋一个权值 $w(e)$, $w(e) = d \cdot k$ 个数据传输或称为流构成了一个集合 $K = \{K_1, K_2, \dots, K_k\}$, $K_i = (s_i, t_i, d_i)$, s_i 和 t_i 分别表示源节点和目的节点, $s_i \neq t_i$ (可以有多条流具有相同的源或目的), d_i 表示流 i 的总的带宽需求。对于流 i , 从 s_i 到 t_i 的路由是一条路径 p_i 。本文使用的是微片缓存的流控,所提出的结构也可用于其他流控机制。

2.3 算法步骤

算法执行过程描述如下:

1) 对于一个给定的应用,由基于遗传算法的映射技术,找出 IP 核到网络节点的近最优映射,并画出映射后的拓扑结构图;

2) 根据网络中节点对通信信息和通信业务的并发信息,画出流图 G ;

3) 由流图 G 生成带权邻接矩阵 M_{cost} ;

4) 在带权邻接矩阵 M_{cost} 中运行弗洛伊德(Floyd)算法,为有通信需求的每对节点生成最短路径。每对节点的最短路径即为 G 中每条流的最短路由。

一个算法执行过程的例子如图 2 所示。算法的时间复杂度与弗洛伊德的算法的时间复杂度相同,为 $O(n^3)$, n 为拓扑结构图中的节点总数。

2.4 虚信道静态分配

BARA 设网络中的每段物理链路含有两条虚信道,分别为 VC0 和 VC1。虚信道静态分配即是为每条流的每段链路

指定一条虚信道。为每条流静态分配虚信道的方式简化了路由器的虚信道分配阶段的设计。因为每条流的每段链路的虚信道号是通过路由算法指定的,而非由仲裁器动态分配^[10]。此外,虚信道静态分配通过将流中的各段链路按路由方向分为两个集合,避免了环路等待,从而保证 BARA 是无死锁的。

定义 2 信道相关图 (Channel Dependence Graph, CDG) $D(V', E')$ 可从流网络 G 获得。方法如下:以 G 中的每条边作为 CDG 中的一个节点。对于 CDG 中的两个节点 $V_1 \in V'$ 与 $V_2 \in V'$,如果分组在 G 中占有与 V_1 相关的边(设为信道 c_i)同时请求与 V_2 相关的边(设为信道 c_j),则称 c_i 与 c_j 存在信道相关,将 V_1 与 V_2 连边。

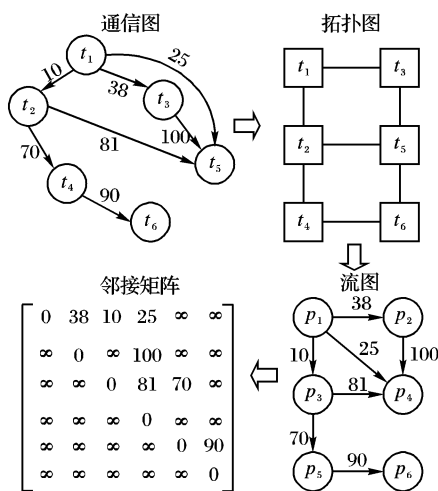


图2 算法执行过程举例

引理 1 如果根据路由函数得到的信道相关图中没有形成环,则可以认为该路由函数是无死锁的^[11]。

下面证明 BARA 是无死锁的。

根据引理 1,需要证明由 BARA 得出的信道相关图是无环的。首先,每段链路,即两个相邻节点间有两条虚信道, VC0 和 VC1。从一个节点出发的分组可以有东、南、西、北四

个输出方向,设东方向为 X^+ 方向,北方向为 Y^+ 方向,西与南方向分别为 X^- 与 Y^- 方向。如图 3(a) 所示。规定若分组从一个节点出发到达它的 X^+/Y^+ 方向邻节点,则这段链路使用 VC0;若分组从一个节点出发到达它的 X^-/Y^- 方向邻节点,则这段链路使用 VC1。

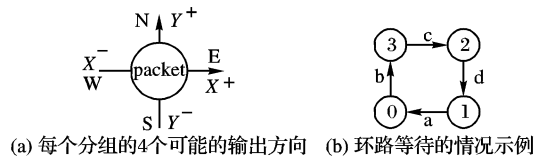


图3 输出方向和环路等待

信道相关图出现环路等待的情形如图 3(b) 所示。其中: 0, 1, 2, 3 分别代表不同的分组; a, b, c, d 代表虚信道。

a, b, c, d 使用的虚信道有三种可能情况。

1) 若 a, b, c, d 均使用 VC0, 按照前面对虚信道分配的规定, 分组应该向 X^+/Y^+ 方向路由。也就是说, 禁止向西和向南的路由方向。不存在由 d 到 c 的分组以及由 c 到 d 的分组。因此, 这种情况下没有环路形成。

2) 若 a, b, c, d 中的部分使用 VC0, 其他使用 VC1。例如, b, c, d 使用 VC0, a 使用 VC1。则 a 只能用于传输 VC1 中的分组, 而不能传输 VC0 中的分组。因此, 这种情况下也不会形成环。

3) 若 a, b, c, d 均使用 VC1, 类似第 1) 种情况, 没有环路形成。

因此, BARA 可以避免环路等待, 它是无死锁的。

2.5 路由表压缩

BARA 采用基于表实现的路由器。随着 NoC 面积的增大, 存储大量目的地址的表需占用大量的存储空间, 路由器存在大量硬件和面积开销。

文献[12]提出了一种路由表压缩的方法, 将路由表按颜色分成不同的群, 实现对路由表的压缩。该压缩方法的基本思想如图 4 所示。

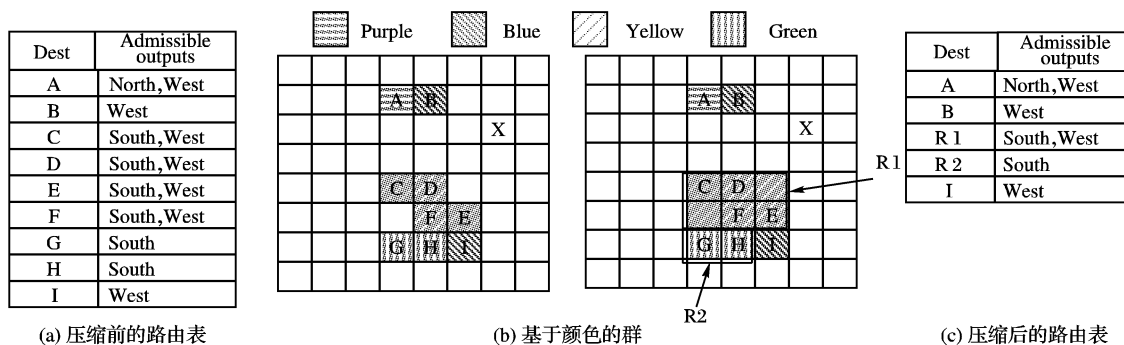


图4 路由表压缩方法示意图

假定有一张与节点 X 的东向输入端口相关联的路由表。考虑到最短路由, 东向的输出端口输出的目的节点应在 NoC 的左边。将允许的输出口方向分别与一种颜色相关联。例如, 向北输出为红色, 向南输出为绿色, 向西输出为蓝色, 向西/北输出为紫色, 向南/西输出为黄色。根据路由表中存储的允许输出的信息, 将所有目的节点分别标记上颜色。将需要经过相同的输出信道转发才能到达的目的节点标记为同一种颜色。由此形成几个基于颜色的群。如图 4(b) 中的几个矩形区域。压缩技术的基本思想是: 无需存储所有的目的节点的

集合, 只需存储群的集合, 如图 4(c) 所示。

使用该压缩技术实现的路由表的造价只占整个路由器价格的一小部分。

3 实验结果与分析

为了验证本文提出的路由技术, 用 C++ 构建模拟器, 该模拟器基于 NoC 模拟器 Nirgam 平台, 实现 4×4 2-D mesh 网络。仿真环境中数据的最小单元是 Flit, 交换机制采用虫孔交换机制。每段链路含有两条虚信道, 每条虚信道的缓存大

小为16个Flit。仿真频率为1GHz,分组之间的时间间隔由数据注入率确定。实验过程从H.263解码过程的通信图^[14]出发,考虑两类映射后的流量模型。一类是H.263解码过程通过随机映射后提取的流量分布模型,另一类是H.263解码过程通过基于遗传算法的映射技术映射后提取的流量分布模型。在两种流量分布模型下比较三种路由算法XY,OE,BARA的时延性能。

3.1 随机映射下的流量评测

本文实现了一个真实的H.263解码过程系统映射,将该应用通过随机映射方法映射到 4×4 2-D mesh拓扑图,分析获得该应用映射后的流量特征。

如图5所示为采用本文算法与维序XY路由算法以及奇偶转弯路由算法(Odd-Even turn routing, OE)^[15]用于该随机映射上获得的结果比较。

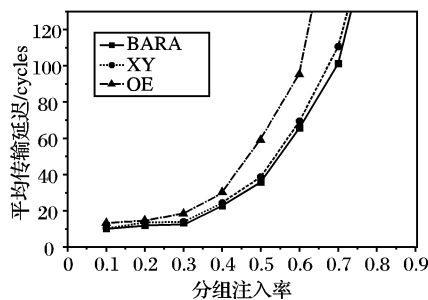


图5 随机映射下的平均时延

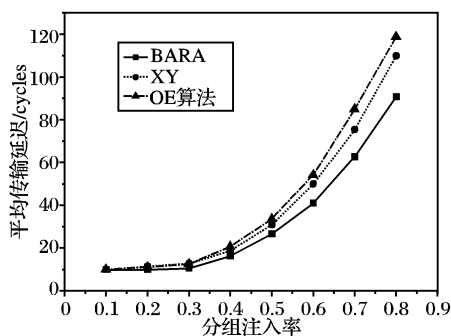


图6 基于遗传算法的映射下的平均时延

3.2 采用基于遗传算法的映射技术映射后的流量评测

采用前面介绍的基于遗传算法的映射技术将H.263解码过程的通信图映射到 4×4 2-D mesh上,分析获得映射后的流量特征。

如图6所示为BARA与XY,OE算法用于该映射上获得的结果比较。仿真结果表明,BARA相比其他两种算法具有更低的延迟,且在该映射下的性能优于随机映射下的性能。

上述实验表明,文中给出的带宽感知的路由算法在不同的流量模型下具有更好的时延性能。这是因为BART充分利用了可用的带宽,减轻了链路负载,减少堵塞。虚信道静态分配提高了流仲裁的公平性^[10],避免一条流堵塞该路径上的其他流。将基于遗传算法的映射技术与带宽感知的路由算法相结合实现了路由算法的二次优化,表现出更好的性能。

4 结语

对于资源受限的NoC系统,为了减轻系统实现开销并优化性能,可以根据设计时已知的应用的信息来设计路由算法。本文给出一种面向应用的路由技术,结合了基于遗传算法的

映射技术与带宽感知的路由算法。在满足设计目标和限制条件的前提下,最优化网络性能。

仿真实验结果表明,对于可预知通信要求的面向应用的NoC,BART可有效降低NoC的数据分组平均传输延迟。该技术可应用于任意拓扑结构的网络模型中。后续的研究工作将考虑把该技术应用到带有故障的拓扑结构中。此外,需要获得更多的关于具体应用的信息,如数据传输的信息等,有针对性地进行路由设计。

参考文献:

- [1] HEMANI A, JANTSCH A, KUMAR S, *et al.* Network on chip: An architecture for billion transistor era[EB/OL]. [2010-06-01]. http://ihome.ust.hk/~ldcse/CSIT560_papers/NOC/norchip-noc.pdf.
- [2] BENINI L, MICHELI G D. Networks on chips: A new SoC paradigm[J]. *Computer*, 2002, 35(1): 70-78.
- [3] DALLY W J, TOWLES B. Route packets, not wires: on-chip interconnection networks[C]// *Proceedings of the Design Automation Conference*. New York: ACM, 2001: 684-689.
- [4] KUMAR S, JANTSCH A, SOININEN J-P, *et al.* A network on chip architecture and design methodology[C]// *Proceedings of the IEEE Computer Society Annual Symposium on VLSI*. Washington, DC: IEEE Computer Society, 2002: 117.
- [5] HELMIG J. Developing core software technologies for TI's OMAPTM platform[EB/OL]. [2010-05-01]. <http://focus.ti.com/pdfs/vf/wireless/omapsoftwaretechnologies.pdf>.
- [6] BENINI L. Application specific NoC design[EB/OL]. [2010-05-01]. http://www.date-conference.com/proceedings/PAPERS/2006/DATE06/PDFFILES/04G_3.PDF.
- [7] PALESI M, HOLSMARK R, KUMAR S, *et al.* Application specific routing algorithms for networks on chip[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2009, 20(3): 316-330.
- [8] TORNERO R O, ORDUA J M, PALESI M, *et al.* A communication-aware topological mapping technique for NoCs[C]// *Proceedings of the 14th International Euro-Par Conference on Parallel Processing*, LNCS 5168. Berlin: Springer-Verlag, 2008: 910-919.
- [9] KINSY M A, CHO M H, WEN T, *et al.* Application-aware deadlock-free oblivious routing[C]// *Proceedings of the International Symposium on Computer Architecture*. New York: ACM Press, 2009: 208-219.
- [10] SHIM K S, CHO M H, KINSY M A, *et al.* Static virtual channel allocation in oblivious routing[EB/OL]. [2010-05-01]. <http://people.csail.mit.edu/devadas/pubs/static-nocs.pdf>.
- [11] DALLY W J, SEITZ C L. Deadlock-free message routing in multiprocessor interconnection networks[J]. *IEEE Transactions on Computers*, 1987, 36(5): 547-553.
- [12] PALESI M, KUMAR S, HOLSMARK R. A method for router table compression for application specific routing in mesh topology NoC architectures[C]// *Proceedings of 6th International Workshop, SAMOS 2006*. Berlin: Springer, 2006: 373-384.
- [13] 张立鹏, 顾华玺, 王长山. 一个新的模块化NoC路由器设计[J]. *计算机应用*, 2009, 29(2): 511-513.
- [14] DICK R. Embedded system synthesis benchmarks suites (E3S). [EB/OL]. [2010-05-01]. <http://www.ece.northwestern.edu/~dickrp/e3s>.
- [15] CHIU G M. The odd-even turn model for adaptive routing[C]// *IEEE Transactions on Parallel and Distributed Systems*, 2000, 11(7): 729-738.