

基于“包-电路交换”的片上网络回退转向路由算法

李丽^{*①} 万健^① 王佳文^① 潘红兵^① 许俊^① 孙敏敏^① 侯宁^②

^①(南京大学微电子设计研究所江苏省光电信息功能材料重点实验室 南京 210093)

^②(合肥工业大学微电子设计研究所 合肥 230009)

摘要: 采用“包-电路交换”的片上路由器,链路的建立通过发送请求包完成,而数据的传输则采用电路形式。传统的路由算法已经不能很好地适应基于“包-电路交换”的片上网络(NoC)新特性。该文根据“包-电路交换”的NoC的特点,提出了一种新的路由算法——回退转向(RT)路由算法,以改善NoC性能。实验结果表明,与动态XY路由算法相比,回退转向路由算法使得网络平均吞吐量和平均包延迟最大分别改善26.7%和11.6%。

关键词: 片上网络(NoC); 路由算法; “包-电路交换”

中图分类号: TP302

文献标识码: A

文章编号: 1009-5896(2011)11-2759-05

DOI: 10.3724/SP.J.1146.2011.00284

NoC Retrograde-turn Routing Algorithm Based on Packet-circuit Switching

Li Li^① Wan Jian^① Wang Jia-wen^① Pan Hong-bing^①

Xu Jun^① Sun Min-min^① Hou Ning^②

^①(Key Laboratory of Advanced Photonic & Electronic Materials, Institute of VLSI Design,
Nanjing University, Nanjing 210093, China)

^②(Institute of VLSI Design, Hefei University of Technology, Hefei 230009, China)

Abstract: The routers on chip, which adopt packet-connection circuit switching, establish the links by sending a request packet and transfer data by circuit switching. Conventional routing algorithms are not suitable for the new features of Network on Chip (NoC) system based on packet-circuit switching. According to these new features, this paper proposes a new routing algorithm, namely Retrograde-Turn (RT) routing algorithm, to improve the performance of the NoC network. Compared with the dynamic XY routing algorithm, the experiment results demonstrate that the new routing algorithm can improve the average throughput and the average latency by 26.7% and 11.6% at best, respectively.

Key words: Network on Chip (NoC); Routing algorithm; Packet-connection circuit switching

1 引言

片上网络(Network on Chip, NoC)是1999年前后几个研究小组提出的一种全新集成电路体系结构^[1-4]。相对于传统的总线架构, NoC在可扩展性、可重用性、设计效率、带宽等方面具有优势,因此成为片上通信领域备受瞩目的话题^[5,6]。近年来,众研究小组已经在NoC领域展开了广泛而深入的研究,如映射算法^[7],路由算法^[8,9]等。路由算法是影响NoC系统性能的重要因素之一。目前,广泛使用的路由算法有静态XY路由算法、动态XY路由算

法(Dy_XY)等。对于不同拓扑结构及不同交换方式的NoC,必将涉及不同路由算法,而这将对其性能产生影响。因此,路由算法的选择是NoC设计中要面临的一个关键问题。

“包-电路交换”是一种适合大通信量的交换方式,其特点是链路的建立通过发送请求包完成,而数据的传输则采用电路形式。在“包-电路交换”中,建立链路时的路由延时与网络状况有关;链路建立以后,数据传输便不依赖于网络状况,数据传输延时小而且可预测,这使得“包-电路交换”可以满足大量、连续数据传输的强实时性要求。常用于虫孔、存储转发、虚拟直通等交换方式的路由算法大多适合通信量不大的应用,而没有考虑到“包-电路交换”的通信特点,因此很难充分发掘基于“包-电路交换”的NoC的潜力。一方面,链路建立是否成功对基于

2011-03-27收到, 2011-09-05改回

国家863计划项目(2008AA01Z135), 国家自然科学基金(60876017)

和江苏省科技支撑计划(BE2009143)资助课题

*通信作者: 李丽 lili@nju.edu.cn

“包-电路交换”NoC 的性能有很大影响;另一方面,随着片上集成度的提高, NoC 中有大量路由节点分布,所以,基于“包-电路交换” NoC 的路由算法应该:

(1)选择合适路径,尽可能成功建立链路,最大化网络通信效率;

(2)复杂度低,实现成本低,减少硬件资源消耗。

根据该思想,本文提出了基于“包-电路交换”的 NoC 回退转向路由算法,以适应“包-电路交换”的特性,提高通信效率。

2 系统平台

2.1 NoC 系统结构

本文基于 2 维网格结构的 NoC 系统级仿真平台开展研究,该平台是由资源节点(resource)、网络接口(Network Interface, NI)、片上路由器(router)和链路(channel)组成。资源节点可以是处理单元、存储单元或片上系统(System on Chip, SoC)。路由器构成了 NoC 网络的通信节点,是其核心部件。4×4 规模的 NoC 网络结构如图 1 所示。

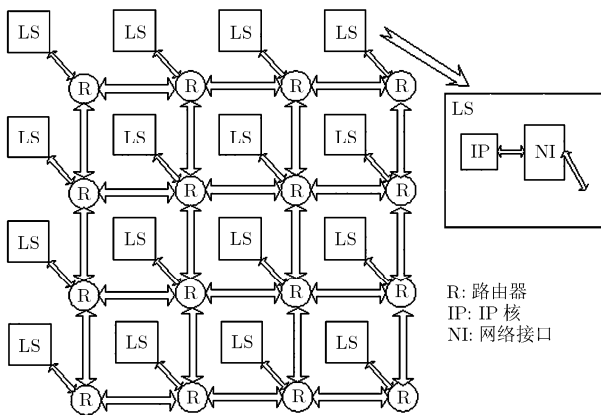


图 1 NoC 系统结构

2.2 路由器

本文所用路由器主要承担 NoC 网络资源节点间的通信任务。每一个路由器节点拥有 5 个双向端口:东(east)、南(south)、西(west)、北(north)和本地(local)。每个端口具有输入和输出通道,如图 2 所示。本地端口负责路由器节点和本地子系统的通信,其它端口负责路由器节点与近邻路由器节点的通信。

3 回退转向路由算法

3.1 动因与简介

在 NoC 系统中,资源节点产生数据包,并将这些数据包交由路由器节点发送到目的资源节点。这

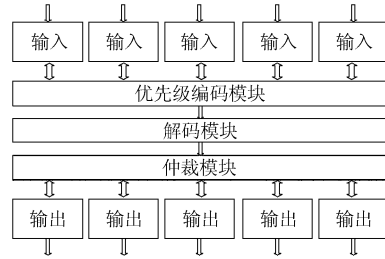


图 2 片上路由器结构

种基于数据包的大量通信任务是并行执行的。因此,当存在多个输入通道请求同一个输出通道进行数据传输时,就会出现拥塞。传统的路由算法在遇到拥塞时,通常会选择等待或是退回到源节点重新请求链路。

基于“包-电路交换”的 NoC 系统在传输数据时,首先由源节点产生链路建立包(包含目的节点网络地址),并将此建立包交给与其相连的路由器节点;路由器节点根据网络情况将建立包发送至目的节点;然后,目的节点反馈一个确认信号至源节点,表示链路建立成功;此后,该链路锁定,源节点与目的节点间通过类似电路交换的方式传输数据。

根据“包-电路交换” NoC 系统的特点,采用传统的路由算法很难发挥其通信潜力。因此,本文提出了回退转向路由算法,全面考虑了网络拥塞情况下的路由选择,提升了链路建立的成功率,从而提高了 NoC 网络的通信性能。

采用回退转向路由算法的路由器节点在处理输入通道的请求时,首先由地址解码模块(decoder)从建立包中读取目的节点的地址,与本地地址进行比较(如图 3 所示),选择可能的输出端口,并将可能的输出端口交给仲裁器模块(arbiter)进行仲裁;仲裁器模块使用 dest 状态寄存器记录各个输入端口可选择的输出端口,并根据一定的优先级顺序和端口占用情况,将输入端口与适当的输出端口相连。某一时刻,一个输出端口只能与一个输入端口相连。当数据传输结束后,被占用的输出端口被释放,此后,其他的输入端口才可以选择该输出端口。如果某输入端口的所有可能的输出端口都被占用,那么该路由器节点将向上游路由器节点发送路由失败信号(fail)。上游路由器节点收到路由失败信号后,查看寄存器 dest 状态,选择其他可能的输出端口,若没有其他可能的输出端口,则此路由器节点向其上游路由器节点发送路由失败信号。依此类推,直到选择合适的路径到达目的节点。

3.2 优势分析

根据上节所述,回退转向路由算法作为一种自

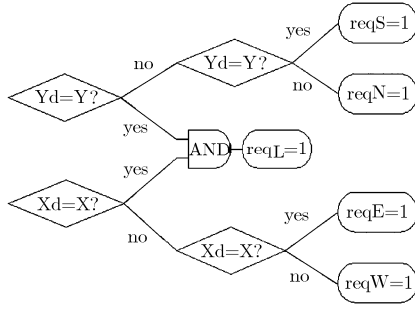


图 3 地址比较机制

适应路由算法，能够根据网络资源的占用情况动态改变路由路径，可以充分利用网络资源；回退转向路由算法在选择路由路径时最多有两个可能的方向，并且不向 180° 方向折回路由，因此所选路径是源节点与目的节点间最短的路径，而且不会引起死锁或活锁的问题。

假设一个 $U \times V$ 的 2 维 mesh 网络，其路由器节点 i 的网络地址为 (i_x, i_y) 。显然，通过路由器节点 i 的建立包有 3 种情况：(1)与路由器节点 i 相连的资源节点 i 发出的建立包；(2)目的节点为资源节点 i 的建立包；(3)需要经过路由器节点 i 传到其他路由器节点的建立包。文献[10]讨论了一个数据包通过路由器节点的概率：

$$P_{j-i} = \begin{cases} 0, & i_y \notin (j_y, d_y) \text{ (or } [d_y, j_y]) \\ & \text{or } i_x \notin (j_x, d_x) \text{ (or } [d_x, j_x]) \\ 1, & i_x = j_x = d_x, i_y \in (j_y, d_y) \text{ (or } [d_y, j_y]) \\ & \text{or } i_y = j_y = d_y, i_x \in (j_x, d_x) \text{ (or } [d_x, j_x]) \\ p, & \text{其他} \end{cases} \quad (1)$$

其中路由器节点 j 的网络地址为 (j_x, j_y) ，目的节点 d 的网络地址为 (d_x, d_y) ， P_{j-i} 表示路由器节点 j 将建立包(目的节点为 d)转发给相邻路由器节点 i 的概率。假设路由器节点 j 处等待转发的建立包的目的地节点地址为 $(d_x > j_x, d_y > j_y)$ ，那么路由器节点 j 可能将建立包转发给路由器节点 $i(i_x = j_x, i_y = j_y + 1)$ 或者转发给路由器节点 $k(k_x = j_x + 1, k_y = j_y)$ 。

对于动态 XY 路由算法，假设每条链路拥塞的概率为 $p(0 < p < 1)$ ，那么经由某个路由节点不能通过其下一级路由节点传递数据包的概率为

$$P_{Dy_XY} = pp + 2(1 - p)pp \quad (2)$$

然而，对于回退转向路由算法，如果路由器节点 j 将建立包转发给路由器节点 k ，但是由于网络拥塞，路由器节点 k 没有可供选择的路由路径，那么根据回退转向路由算法，路由器节点 j 仍有可能将

建立包转发给路由器节点 i ，因此，通过路由器节点 i 的概率增加，而经由某个路由节点不能通过其下一级路由节点传递数据包的概率减小为

$$P_{RT} = pp + 2(1 - p)ppp[1 + p(1 - p)] \quad (3)$$

由于 p 在 0 到 1 的范围内，所以 P_{RT} 始终小于 P_{Dy_XY} ，由此可知，与其他自适应路由算法相比，回退转向路由算法增加了链路建立成功的概率，从而能够更好地避免拥塞，拥有更高的平均吞吐量和更低平均包延迟。

3.3 伪码

在应用回退转向路由算法的路由器节点中，主要通过函数 `retro_granting()` 实现回退转向路由算法。其伪码如表 1 所示。

表 1 RT 路由算法伪码

```

int retro_granting()
{
    int i = 0, j = 0, k = 0;
    int in_flag, out_flag;
    /* in_flag 记录输入端口号, out_flag 记录输出端口号 */
    int new_out_flag [ N_PORT ];
    /* new_out_flag 队列记录各输入端口重新选择的输出端口号 */
    for(i = 0; i < N_PORT; i++)
        {if ( fail [ i ] == 1 )
            /* 判断下游路由器节点是否返回有效的 fail 信号 */
            {out_flag = i;
            for (j = 0; j < N_PORT; j++)
                {//找出对应的输入端口号
                if ( Conn [ j ][ out_flag ] == 1 )
                    /* Conn 寄存器中记录了输入端口与输出端口的对应关系 */
                    in_flag = j;
                }
            }
        }
    for (k = 0; k < N_PORT; k++)
        if ( Dest [ in_flag ][ k ] == 1 && k != out_flag )
            new_out_flag [ in_flag ] = k;
    }
    return( new_out_flag )
}
    
```

3.4 实现

回退转向路由算法在片上路由器中实现时由 4 个部分组成：RETRO 模块，DEST 模块，CONN 模块和 OCU 模块。结构如图 4 所示。RETRO 模块执行路由策略，接收下游路由器节点反馈的路由失败信号 `fail[i]`；DEST 模块接收地址解码后满足路由条件的可能的输出端口号 `req_in[i]` 信号，结合输出

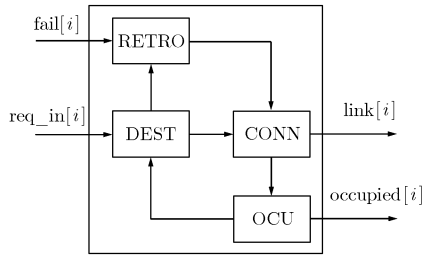


图 4 算法实现结构

通道占用情况记录这些可能的输出通道号；CONN 模块产生输入端口与输出端口连接信号 $link[i]$ ；OCU 模块产生输出端口占用信号 $occupied[i]$ 。

4 实验与分析

4.1 实验性能指标

对于 NoC 网络系统，评价其性能的主要指标是平均吞吐量(throughput)和平均包延迟(latency)^[11]。

其定义如下：平均吞吐量 = $\frac{\text{包数量}}{\text{节点数} \times \text{模拟时间}}$ ，平均包延迟 = $\frac{\text{包传输延迟总和}}{\text{包数量}}$ 。

4.2 实验条件

本文所用的实验平台是基于“包-电路交换”的 $4 \times 4 \times 2$ 维 mesh 结构 NoC 仿真平台。系统设置：“包-电路交换”，回退转向路由算法，数据片宽度为 34 bit，包长度为 50 个数据片。此处将包长度设置为 50 个数据片主要是因为“包-电路交换”的特点就是适合大数据包的传输。系统每次模拟运行的时间是 25,000 个时钟周期。以上是系统的默认环境设置，如果没有特殊说明均采用此系统设置。此系统中，本地系统负责发送数据和接收数据；通过改变数据包的注入率可以实现不同的网络拥塞情况，从而可以验证回退转向路由算法在不同网络负载情况下的性能。

由于实践中广泛应用动态 XY 路由算法，并且其表现出良好的性能，因此实验中将其作为回退转向(RT)路由算法的比较参考。

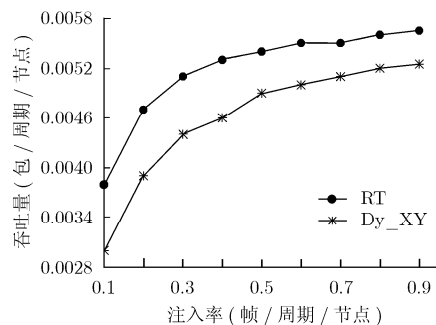


图 5 平均吞吐量与注入率的关系

4.3 实验结果

实验结果如图 5 和图 6 所示。其中，图 5 是平均吞吐量曲线，其横坐标表示数据包注入率，纵坐标表示平均吞吐率；需要特别说明的是，图中注入速率，如 0.1 帧/周期/节点并不是指每个周期每个节点注入 0.1 帧，因为包电路交换机制和普通包交换机制不同，一旦建立链路，传输是连续的，所以这里的含义是节点注入包头帧，即发起一轮传输的概率为 0.1。图 6 是平均包延迟曲线，其横坐标意义与图 5 相同，纵坐标则表示平均包延迟。

实验结果显示，与动态 XY(DY-XY)路由算法相比，回退转向(RT)路由算法在平均吞吐量和平均包延迟上都有很大改善，最大改善程度分别能够达到 26.7%和 11.6%。同时，在不同的网络负载情况下，回退转向路由算法都优于动态 XY 路由算法。可见，实验结果与之前的分析结果一致。

4.4 综合结果

由上节的实验结果可知，与动态 XY 路由算法相比，回退转向路由算法性能优势明显。本节主要说明回退转向路由算法的实现成本代价，仍然以动态 XY 路由算法为参考。本文使用 Verilog 硬件描述语言对以上两种基于不同路由算法的路由器模型进行建模，采用 ISE 10.1.3 工具，在 Xilinx 公司 Virtex 5 系列的 XC5VLX330T 芯片上进行了设计实现。综合结果如表 2 所示。

表 2 路由器模块硬件资源消耗

路由	查找表	寄存器
Dy_XY	1053	352
RT	1142	359

综合结果显示，与动态 XY 路由算法相比，回退转向路由算法查找表(逻辑单元)有少许增加(8.5%)，而寄存器单元用量几乎相同。可见，实现回退转向路由算法导致路由器硬件成本增加很少。

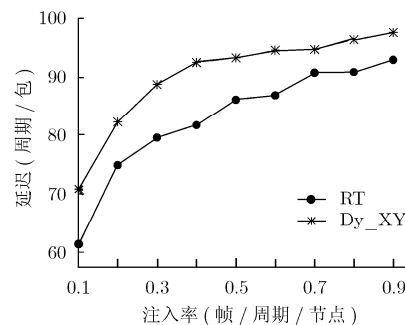


图 6 平均包延迟与注入率的关系

5 结束语

本文提出了一种新的路由算法——回退转向路由算法。该方法充分考虑了 NoC 拥塞和“包-电路交换”的特点，提高了 NoC 资源的利用率和链路建立的成功率，因此它改善了 NoC 的性能。实验结果显示，与动态 XY 路由算法相比，回退转向路由算法在平均吞吐量和平均包延迟性能指标上最大分别能改善 26.7%和 11.6%，而且其优势覆盖了整个负载条件。综合结果表明，回退转向路由算法导致路由器组合逻辑少许增加(8.5%)。由此可见，回退转向路由算法是一种新颖、高效率、低实现成本的 NoC 路由算法，它除了可以很好地适用于包-电路交换的路由方法，对于其他路由机制也有较好的借鉴意义。

参 考 文 献

- [1] Soteriou V, Ramanujam R S, Lin B, *et al.* A high-throughput distributed shared-buffer NoC router[J]. *IEEE Computer Architecture Letters*, 2009, 8(1): 21–24.
- [2] Pande P P, Zhu H B, Ganguly A, *et al.* Energy reduction through crosstalk avoidance coding in NoC paradigm[C]. Proceedings of the 9th EUROMICRO Conference on Digital System Design (DSD'06), Cavtat, HR: IEEE, 2006: 689–695.
- [3] Benini L and Micheli G D. Networks on chips: a new SoC paradigm[J]. *IEEE Computer*, 2002, 35(1): 70–78.
- [4] Ahmad B and Arslan T. Dynamically reconfigurable NoC for reconfigurable MPSoC[C]. IEEE 2005 Custom Integrated Circuits Conference, San Jose, CA: IEEE, 2005: 277–280.
- [5] Marculescu R, Ogras U Y, Li-Shiuan Peh, *et al.* Outstanding research problems in NoC design: system, microarchitecture, and circuit perspectives[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2009, 28(1): 3–21.
- [6] Freitas H C, Santos T G S, and Navaux P O A. Design of programmable NoC router architecture on FPGA for multi-cluster NoCs[J]. *Electronics Letters*, 2008, 44(16): 969–971.
- [7] 陈庚生, 陈亦欧, 胡剑浩. 面向实时数字信号系统关键链路延时的 NoC 映射方法研究[J]. *电子与信息学报*, 2010, 32(7): 1638–1643.
- [8] Chen Geng-sheng, Chen Yi-ou, and Hu Jian-hao. A novel critical delay-aware mapping method for real-time digital signal systems with NoC platform[J]. *Journal of Electronics & Information Technology*, 2010, 32(7): 1638–1643.
- [9] Senin I V, Mhamdi L, and Goossens K. Efficient multicast support in buffered crossbars using Networks on Chip[C]. Global Telecommunications Conference, 2009, Honolulu, HI. IEEE, Nov. 2009: 1–7.
- [10] 王坚, 李玉柏, 蒋勇男. 片上网络通信性能分析建模与缓存分配优化算法[J]. *电子与信息学报*, 2009, 31(5): 1059–1062.
- [11] Wang Jian, Li Yu-bai, and Jiang Yong-nan. Communication performance analytical model and buffer allocation optimizing algorithm for Network-on-Chip[J]. *Journal of Electronics & Information Technology*, 2009, 31(5): 1059–1062.
- [12] Li Ming, Zeng Qing-an, and Jone Wen-ben. DyXY-a proximity congestion-aware deadlock-free dynamic routing method for Network on Chip[C]. DAC 2006, San Francisco, California, USA, July 24–28, 2006: 849–852.
- [13] Yang Sheng guang, Li Li, Xu Yi, *et al.* A power-aware adaptive routing scheme for network on a chip[C]. In 7th International Conference on ASIC Proceedings(ASICON 2007), Guilin, China: IEEE, Oct. 26–29, 2007, 2: 1301–1304.

李 丽： 女，1975 年生，博士，教授，研究方向为多处理器片上系统设计方法、可重构计算和 IP-SoC 设计方法。

万 健： 男，1985 年生，硕士，研究方向为超大规模集成电路设计及片上网络。