

基于路径的无死锁 NoC 多播路由器的设计

张 颖, 吴 宁, 葛 芬

(南京航空航天大学 电子信息工程学院, 江苏 南京 210016)

摘 要: 针对 NoC 中并行处理事务的增多, 将计算机网络中的多播传输模式引入到 NoC 路由器的设计中来, 提出了一种新颖的基于路径的多播路由器设计模式. 采用数据包中带有多个目的地址及数据子包的模式, 实现了数据包到多个目的节点的高效传输, 同时, 通过引入 Hamiltonian 路径的无死锁处理模式, 避免了数据包传输过程中死锁的发生. 采用 Verilog HDL 实现了该模型的 RTL 级设计, 并用 Synopsys EDA 软件完成功能仿真、时序验证及综合.

关键词: 片上网络; 多播; 路由器; 无死锁

中图分类号: TN407

文献标识码: A

文章编号: 1000-7180(2014)11-0001-05

Design of a Deadlock-free Multicasting Router for NoC Based on Path

ZHANG Ying, WU Ning, GE Fen

(College of Electronic and Communication Engineering, Nanjing University of
Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Faced with the increase of concurrent processing for NoC, a novel multicast router based on path is proposed in this paper, which introduces the multicast of internet to the design of NoC routers. This router adopts packets which contains a number of destinations and sub-packets and realizes the efficient transmission of packets to many destinations. At the same time, the router avoids the occurrence of deadlock routing in the process of transmitting packets referenced to the deadlock-free processing mode of Hamiltonian path. The model is designed by VerilogHDL and its function simulation, timing evaluation and synthesis are implemented on Synopsys EDA platform.

Key words: NoC; multicast; router; deadlock-free

1 引言

在片上网络 NoC(Network on Chip)的构成部件中, 路由器是最关键的部分, 直接决定着 NoC 的传输性能和工作效率^[1]. 现有 NoC 路由器设计大都关注于单播传输功能的实现, 但是未来的系统芯片将以 NoC 的互连方式集成成百上千个 IP 核, 并且存在大量的同构核^[2]. 在并行图像处理算法、NoC 内嵌 IP 核的并行测试^[3]等状况下, 核内需要并行处理的数据量巨增, 传统的单播传输模式逐渐不能满足系统需求. 多播传输模式可以有效地缓解 NoC 通

信链路的传输负担, 提升网络带宽的利用率, 因而可作为 NoC 高速并行数据传输的解决方案.

在进行多播传输时, 多播的所有目的地址构成一个多播组, 每个目的地址被称作多播组成员. 实现多播传输的方式主要有以下几种^[4]. (1) 基于单播的多播通信模式: 源节点连续地向每个多播组成员发送数据, 即用单播的形式实现多播通信. 这种方式会浪费网络带宽, 此外, 由于需要源节点拥有完整的多播组成员清单, 从而会增加源端开销. (2) 基于树的多播通信模式: 多播树是覆盖了所有多播组成员的一棵生成树, 它能随着多播组成员的变更动态地自我调整. 源

收稿日期: 2014-03-27; 修回日期: 2014-05-23

基金项目: 国家自然科学基金项目(61376025, 61106018); 国防基础科研项目(B2520110008); 江苏省科技支撑计划(BY2013003-11)

节点只需发送一次数据,然后通过数据在多播树的分支处的复制和转发,从而到达每个目的节点.这种方式硬件实现比较复杂,而且如果一个节点出现堵塞,则后续的节点都出现堵塞.(3)基于路径的多播传输模式:源节点所发送的数据的包头微片中包含了所有多播组成员的节点地址,在传输的过程中,按照某种方式依次访问每个多播组成员,在每个目的节点都拷贝一份数据送给当前的网络接口,最终把数据送到所有的目的地址.基于路径的多播模式中,源节点每次只需发送一次数据,就可以送到每个目的节点,提升了传输效率.本文所设计的多播传输模式就是采用该方法.值得一提的是,在多播传输模式中,数据包由于需要在多个目的节点间传输,很容易产生死锁.本文所设计的多播路由器参考 Hamiltonian 路径^[5]的设计方法,避免路由器模型中死锁的发生.

2 基于路径的多播路由策略

2.1 基于路径的多播路由算法

NoC 采用基于数据包的形式进行数据传输.如图 1 所示,数据包分为包头微片和包负载微片,每个微片长度为 32 位.包头微片中包含了所有多播组成员的节点地址及到达各个目的节点子包的序数,包负载微片包含到多个目的节点的子包.本文所采用的包头为多个微片,如果数据包所传输到的目的节点数多于两个,则在第一个包头微片后面会继续添加目的节点信息,多播地址的数量由第一个包头微片的 4 至 1 位表示.包头的最后一个微片是定义各个子包的长度信息,即每个子包所包含的微片数量.包负载为所要传输的数据,由各个子包构成,每个子包的长度由包头中相关信息确定.

从源节点到各目的节点间的传输采用 XY 维序路由算法^[6].当数据包到达路由器时,根据多播地址数,检查多播成员组中排在最后一个目的地址是否与当前路由器地址相同,如果不同,则继续依据路由算法传输数据;如果与该目的地址相同,则把数据包复制一份传输到当前节点的本地端口,经过解包后输出给相对应的 IP 核.与此同时,数据包中的多播地址数减 1,数据包继续按照 XY 维序路由算法向当前排在最后的目的地地址传输,直到把数据包传送给多播成员组的所有节点.

2.2 无死锁处理

XY 维序路由算法在单播模式下不会产生死锁,但在多播模式下,由于数据包在源节点及各个目的节点中的两两节点间传输时应用该算法,但整个数据包



图 1 多播数据包结构

的传输路径却与原路由算法不再符合,因而可能出现死锁.为此,要对多播模式下的路由算法进行改进以避免死锁的发生.参照 Hamiltonian 路径的模式,把整个网络的路由器进行编号,其顺序如图 2 所示,以源节点为分界线,如图中的路由器 9,整个网络被分为两个子集,一个是比源节点编号大的路由器集合,该子集以升序的模式排列;另一个子集是比源节点小的子集,其中的路由器以降序排列.当需要发送数据包的时候,把多播成员组分为两个部分:比源节点地址大的成员构成一个多播数据包,如图 2 中的 13、15,对这些成员组的访问以升序的模式进行;另外把目的节点比源节点小的成员组构成另一个数据包,如 7、5、1 目的节点,对这些成员的访问以降序的模式进行.以上对每一个目的地址的访问还采用传统 XY 维序路由算法,且每一次传输路径都禁止 180° 转弯.以图 2 为例,源节点为路由器 9,送到目的节点 13、15 的数据构成一个多播数据包,数据包传输路径为 $9 \rightarrow 13 \rightarrow 15$;送到节点 7、5、1 的数据构成另一个多播数据包,数据包传输路径为 $9 \rightarrow 7 \rightarrow 5 \rightarrow 1$.

3 路由单元的设计与实现

3.1 路由器的整体结构

路由器工作在全双工模式,因此东、西、南、北、本地五个方向的传输通道都有输入和输出端口.东、西、南、北四个方面采用相同的电路设计,本地部分模块即网络适配器,用于数据包的组包和解包,是连接 NoC 通信结构和嵌入式 IP 核两部分的接口^[7].如图 3 所示,路由器主要由以下几部分构成:输入校验与缓冲模块、路由模块(Routing)、仲裁模块(Arbitrator)及交换开关等.

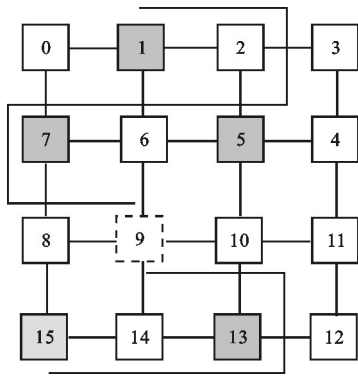


图2 Hamiltonian 路由算法示意图

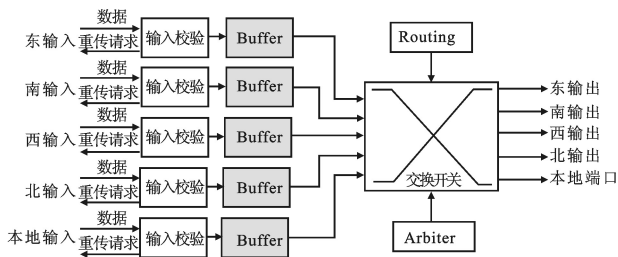


图3 NoC 路由器结构图

3.2 输入校验及缓冲区

根据 NoC 的使用情况,本文所采用的微片长度为 32 位,所有微片的最后一位为偶校验位.为了降低面积及功耗开销,只在每个方向的输入端口设置了缓冲区.图 4 所示为输入校验模块,表 1 所示为其端口说明.当前级数据发送请求信号 Pre_data_req 为高电平时,表示有数据要送入,如果当前缓冲区未滿 ($Ram_full=0$),则向前一级路由器发送同意接收数据的反馈信号,该路由器收到反馈信号后开始传输数据.当有数据进入输入端口时,首先需要进行数据校验,如果 $\neg Pre_data_in=0$,即输入数据符合偶校验,该数据被存入缓冲区;如果不为 0,则 Pre_n_ack

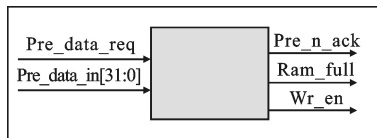


图4 输入校验模块

表1 输入校验模块的端口说明

端口名称	端口类型	端口说明
$Pre_data_in[31:0]$	input	上级路由器传入的数据微片
Pre_data_req	input	上级路由器传入的发送数据请求信号
Pre_n_ack	output	反馈给上级路由器的重传请求信号,高电平有效
Ram_full	output	缓冲区满信号,高电平有效
Wr_en	output	缓冲区写使能信号,高电平有效

置高,即向前一级路由器发出重传请求,前级路由器重传该数据,直到送入数据符合偶校验.

3.3 多播路由单元

路由单元的作用是根据收到的包头信息对数据包的转发方向做出选择,并向该方向的输出端口发起传输请求,在得到同意使用端口的情况下,把数据包向该端口传输.本文所设计的多播路由单元根据包头信息提取多播成员组成个数 $multi_number$,路由单元每次只检查多播地址中位于第 $multi_number$ 个的目的地址是否与该路由节点地址相同,如果不同,则根据 XY 维序路由算法确定数据包的传输方向,并向该方向的输出端口发起数据传输请求,得到同意后开始传输数据包;如果当前路由器为包头信息中的第 $multi_number$ 个目的地址,则需把备份数据包送入路由器的本地端口,本地端口根据该目的节点所带数据子包的序号提取数据微片,完成解包功能,并把解包后的数据送入 IP 核.其路由算法流程如图 5 所示.

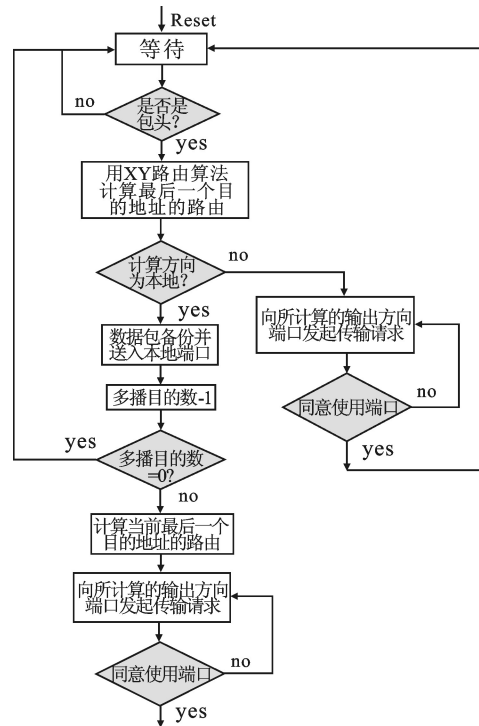


图5 多播路由算法流程图

如果 $multi_number$ 值为 1,则数据包传到当前路由器所连接的 IP 核后整个过程结束,如果 $multi_number > 1$ 则表示还需要把数据包送给 $multi_number - 1$ 个目的地址,此时对 $multi_number$ 减 1,即把当前的第 $multi_number - 1$ 个目的地址作为下一个路由目的地址.在把数据包送入本地之前先向由路由算法确定的目的节点的输出端口发起传输请求,在得到同意使用输出端口的反馈信号后,向该

端口传输数据并备份, 备份数据送给当前路由器的本地端口. 经过多个路由器的传输, 直到把数据包送到多播成员组的每一个目的地址.

3.4 仲裁算法

本文所采用的是基于预防局部拥塞的轮询仲裁算法^[6]. 图 6 所示为本文中使用的仲裁模块, 表 2 所示为其端口说明. 仲裁模块的输入是来自各个方向路由计算模块的传输请求信号 Pre_req_E/S/W/N/L 及各个方向缓冲区剩余状况 Ram_value_E/S/W/N/L, 采用先到先服务的原则. 当有超过一个方向输入端口同时向当前输出端口发起数据传输请求时, 这时就要进行仲裁, 将仲裁结果信号反馈给各输入模块, 获得传输权的信号开始传输数据. 当多个端口同时发起端口申请时, 首先根据各端口反馈过来的缓存区使用情况进行仲裁, 为防止出现局部拥塞, 缓存区剩余空间少的端口优先获得传输权, 如果存在多个端口缓冲区使用情况相同的状况, 则根据东、南、西、北、本地的顺序进行优先权排序. 获得最高优先权的方向先使用本端口传输数据, 当该方向数据包发送完后, 优先权次低的方向开始获得使用端口资格, 直到最低优先权的方向也传输完数据, 然后进入下一轮的仲裁过程.

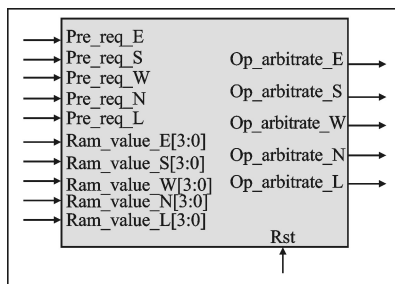


图 6 仲裁模块

3.5 交换开关

交换开关主要实现把路由器输入的数据传输到路由计算模块所选择方向的输出端口. 考虑到 NoC 中资源有限的问题, 本文设计了简单、高效的基于多路复用的 Crossbar Switch. 如图 7 所示, 它把每个

方向的传输端口都连接到交换开关的输入端, 根据每一个多路选择器的控制信号, 对多路数据有选择性输出, 从而实现数据包在路由器之间的传输.

表 2 仲裁模块的端口说明

端口名称	端口类型	端口说明
Rst	input	复位信号, 高电平有效
Pre_req_E	input	东侧端口数据传输请求信号, 高电平有效
...	input	...
Pre_req_L	input	本地端口数据传输请求信号, 高电平有效
Ram_value_E[3:0]	input	本地端口缓冲区剩余空间
...	input	...
Ram_value_E[3:0]	input	本地端口缓冲区剩余空间
Op_arbitrate_E	output	东侧端口申请使用输出端口的仲裁结果
...	output	...
Op_arbitrate_E	output	本地端口申请使用输出端口的仲裁结果

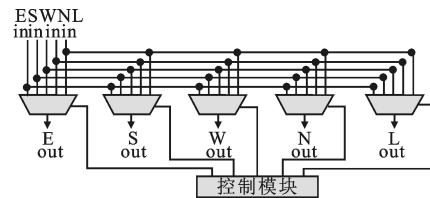


图 7 交换开关结构示意图

4 仿真与结果分析

4.1 功能验证

对于一个 NoC 网络来说, 路由器的优劣直接关系到整个网络的通信状况, 因此对于路由器功能的验证是整个设计流程中至关重要的一环. 路由器通过本地接口(NA)完成与嵌入式 IP 核的连接. 并由 NA 进行组包、解包、包发送、包交付的操作. 在仿真过程中, 路由器 ID 设置为 0101, 从东输入端口输入一多播数据包, 分别包含三个子包, 子包 1 目的地址为 1111, 子包长为 2; 子包 2 目的地址为 1110, 子包长为 3; 子包 3 目的地址为 0101(本地), 子包长为 2, 图 8 为多播路由算法仿真波形图.

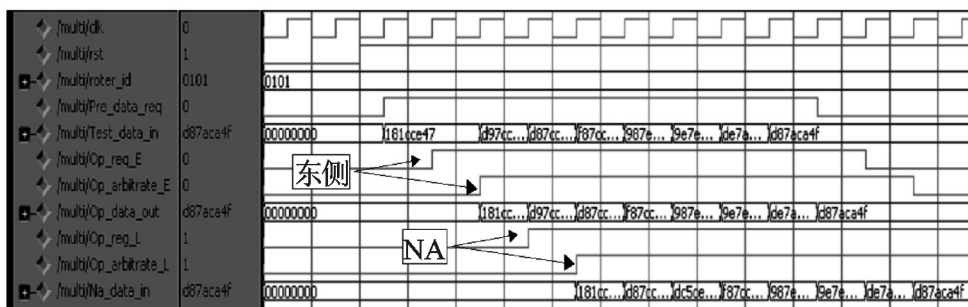


图 8 多播路由算法仿真时序图

根据本文多播算法的设计,首先计算最后一个子包的路由方向,最后一个目的地址为本地,所以进而计算倒数第二个数据子包的路由方向.倒数第二个子包的目的地址是 1110,X、Y 方向都大于本地地址,由于采用的是 XY 路由算法,所以向东侧输出口发起数据传输请求.对 Pre_req_E 信号置位,同时把包头信息中多播地址数减 1.当得到东侧端口的仲裁反馈信号 Op_arbitrate_E 为高电平后,开始向东侧输出口传输数据并将数据备份,然后向本地网络适配器输入端口发起数据传输请求.在得到同意接收数据的信号后,向 NA 发送备份数据.当两个方向的传输结束后,对应方向的传输请求信号变为低电平,本次传输结束.经过多次验证,本文设计的路由器能够正确的实现多播传输.

4.2 路由单元面积功耗开销分析

在保证功能实现的前提下,路由单元的功耗和面积开销也是衡量性能的两个重要指标.路由单元在正常工作时,其消耗的总功耗主要由两部分构成,即器件的静态功耗和动态功耗.静态功耗是芯片在待机状态下所消耗的功耗,主要是由电路的漏电流所产生;动态功耗主要由短路电流和电路内部节点充放电所引起的功耗组成.

功耗和面积的分析采用相同的约束条件,数据都是在对所设计的 RTL 级代码综合后获得.本文的综合是在 SMIC 0.18 μm 工艺库下进行的,由于适配单元、本地 IP 核和路由单元使用的是不同的时钟域,所以路由单元和适配单元分开综合,路由器的面积开销数据如表 3 所示.

表 3 支持多播传输的路由器面积开销

	组合电路 面积/ μm^2	时序电路 面积/ μm^2	连线面积/ μm^2	总面积/ μm^2	所占 百分比/%
路由单元	99 043.69	125 936.44	1 527.00	226 507.13	64.33
适配单元	41 173.12	83 663.79	736.15	125 573.06	35.67

综合后的面积分为组合电路面积、时序电路面积及连线面积,其中,路由单元面积占整个路由器面积的 64.33%,网络适配单元面积占总面积的 35.67%.造成这一现象的原因在于路由单元包括了四个方向的输入输出端口,每个输入端口都有数据缓存单元,而网络适配单元所使用的缓存空间相对较小,所以占用面积较小.

本文所采用的通信网络时钟频率为 200 MHz,网络适配器和 IP 核的工作频率都是 50 MHz,供电电压为 1.8 V,路由器的功耗开销如表 4 所示.

表 4 支持多播传输的路由器功耗开销

	翻转功耗/ mW	短路功耗/ mW	漏电功耗/ μW	总功耗/ mW	所占 百分比/%
路由单元	9.463	4.086	12.54	13.561 5	65.17
适配单元	5.423	1.823	2.11	7.2481	34.83

路由器功耗开销获得与面积开销相似的百分比,其中路由单元的功耗开销占总开销的 65.17%,适配单元的功耗开销占总功耗的 34.83%.由此可以看出,缓存单元不仅是面积开销的主要部件,同时也是功耗开销的主要部件.

5 结束语

本文所设计并实现的无死锁多播 NoC 路由器采用基于路径的路由算法,在设计过程中,参考了 Hamiltonian 路径避免死锁的处理模式.仿真结果表明,该设计可以实现多播模式的传输和避免死锁的发生.相对于传统的单播传输路由器,该路由器具有更高的并行传输能力.此外,在具有大量同构核的 NoC 系统中,由于多个同构核只需在数据包中保存一个数据子包,所以有效地降低了网络中的数据传输量,提升了传输效率,为高效 NoC 通信模式的实现提供了有益的解决方案.

参考文献:

- [1] 周帅,李丽,潘红兵,等.基于微码控制器的 NoC 网络接口设计[J].微电子学与计算机,2012,29(12):7-11.
- [2] Benini L, De D, Micheli. Networks on chips: a new SoC paradigm [J]. IEEE Computer, 2002, 35 (1): 70-78.
- [3] 余燕平,赵问道,孟利民,等.基于关键节点和多播节点的多播路由算法[J].电路与系统学报,2003,8(2):43-47.
- [4] 刘华.片上网络多播通信关键技术研究[D].武汉:武汉理工大学,2011.
- [5] Ran Libeskind-Hadas, Kevin Watkins, Thomas Hehre. Fault-tolerant multicast routing in the mesh with no virtual channels[C]// Proceedings of Second International Symposium on High-Performance Computer Architecture. Shanghai, IEEE, 1996:180-190.
- [6] 万玉鹏,吴宁.基于自适应算法的 NoC 路由单元的系统设计[J].微电子学与计算机,2010,27(5):42-45.
- [7] 吴飞.片上网络适配单元的设计与实现[D].南京:南京航空航天大学,2007.

作者简介:

张颖女,(1977-),博士研究生,讲师.研究方向为数字系统设计与计算机应用、集成电路设计与测试.