适用于虚通道路由器的高性能 round-robin仲裁器

张 哲 高小鹏 龙 翔

(北京航空航天大学 计算机学院, 北京 100083)

摘 要: 分析了交换结构规模不同的2类典型的虚通道路由器体系结构, 指出了交换结构规模对调度机制和仲裁器体系结构的显著影响. 在轮转(round-robin)仲裁器PPE的基础上, 提出了分别适用于上述2类路由器的改进型PPE仲裁器体系结构. 改进型仲裁器将原有关键路径中的处理步骤转移到非关键路径中, 并行实现了关键路径中的重要处理步骤, 通过缩短关键路径改善了仲裁器的性能; 同时通过调整仲裁流程消除了上述改进对资源成本的负面影响. 基于TSMC 0.35μm CMOS工艺实现了不同规模的PPE和2种改进型仲裁器体系结构. 实验结果表明上述2种改进型仲裁器比PPE仲裁器的性能分别提高12%和17%以上, 并且在多数情况下同时减少了实现面积.

关 键 词: 虚通道; 路由器; 调度; 仲裁器; 轮转

中图分类号: TP393

文献标识码: A 文章编号: 1001-5965(2007)06-0743-05

High performance round-robin arbiter for virtual channel router

Zhang Zhe Gao Xiaopeng Long Xiang

(School of Computer Science and Technology, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract Two typical virtual-channel router architectures were analyzed, whose switch fabrics were in different scales. The scale of switch fabric was indicated to lay great impact on scheduler and arbiters in the router. Two arbiter architectures were proposed as variations of PPE, a round-robin arbiter. By moving some steps on the critical paths to the non-critical paths and having some critical steps parallelized, the proposed variations shortened the critical paths over PPE to improve the performance. Some steps were adjusted, so that the side-effect on area caused by above modifications was eliminated. All arbiters including PPE and the proposed variations were evaluated in several scales, using TSMC 0.35μm CMOS technology. The results show that the proposed ones improve the performance by more than 12% and 17% over PPE respectively, and also have less area cost in most cases.

Key words: virtual channel; routers; scheduling; arbiter; round-robin

虚通道(virtual channel)技术是一种被广泛应用于并行计算机网络中的流控技术, 采用该技术的路由器称为虚通道路由器. 虚通道路由器通过为每个缓冲报文的端口提供一组缓冲区(即虚通道), 消除物理信道与缓冲区的耦合, 有效地改善了网络的吞吐率[1]. 在最常见的采用输入排队(IQ, Input Queueing)和分布式调度机制的虚通道路由器中, 仲裁器是决定调度性能的重要因素. 目前的虚通道路由器通常采用以PPE[2]为代表的轮转(round-robin)仲裁器.

虚通道路由器的体系结构对分布式调度机制和仲裁器体系结构的设计具有很大影响, 其中交换结构是最突出的影响因素. 本文分析了2种交换结构规模不同的虚通道路由器对调度机制的需求, 并指出了轮转仲裁器PPE在2种路由器中的性能不足. 在此基础上提出了分别适用于这2种路由器的改进型PPE仲裁器体系结构. 实验结果表明, 二者不仅性能分别比PPE提高了12%和

收稿日期: 2006-08-07

作者简介: 张 哲(1978-), 男, 河北深县人, 博士生, zzhang@cse.buaa.edu.cn.

17%以上,而且在多数情况下还有效地降低了资源成本.

1 虚通道路由器的体系结构和调度机制

如图 1所示,虚通道路由器可以采用不同规模的交换结构,为每个物理通道或者虚通道提供独立的转发端口.二者的差别在于虚通道对交换结构端口的访问是否存在冲突.本文分别称之为有冲突虚通道路由器和无冲突虚通道路由器,并且以图 1中的采用输入排队的 $N \times N$ 路由器为基础进行讨论.

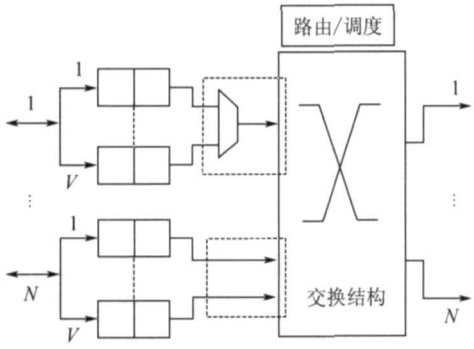

图 1 虚通道路由器的体系结构

在有冲突虚通道路由器中,调度机制需要解决来自同一组虚通道的报文对交换结构的输入端口的竞争,以及所有目的端口相同的报文对同一交换结构的输出端口的竞争(即外部阻塞).以 SLIP[3]为代表的分布式调度机制具有良好的并行性和可扩展性,目前得到了最为广泛的应用[4].分布式调度机制为每个输入端口和输出端口提供独立的仲裁器,通过迭代仲裁达到交换结构的极大利用.每次迭代过程包括请求(request)、响应(grant)和接受(accept)3个步骤.

无冲突虚通道路由器中的交换结构规模增大为前者的 V 倍,消除了关于交换结构的输入端口的竞争,提高了路由器的吞吐率.由于其中只存在外部阻塞,因此,路由器只需为每个输出端口提供独立的仲裁器,一次仲裁即可达到资源的最大利用.

2 PPE仲裁器及性能分析

虚通道路由器的体系结构对调度机制的影响非常显著,同时也对实现调度机制的仲裁器产生了影响.本节首先描述了 PPE 的体系结构,然后分析其在上述 2种路由器中的性能.

2 1 PPE体系结构

PPE采用了集中式体系结构,比采用树形结构的其它轮转仲裁器具有更好的公平性[5-7],在各种输入排队路由器设计中得到了广泛的应用.

图 2描述了 PPE 的体系结构. Req和 Gnt分别是请求信号和仲裁结果信号; P_enc指示了具有最高优先级的请求在 Req中的位置,其它请求信号的优先级自该位置向高位方向循环递减; Encode是 N: $\log_2 N$ 编码器; Smpl_PE是固定优先级仲裁器; ToThermo是温度计编码器(Thermoter Encoder).

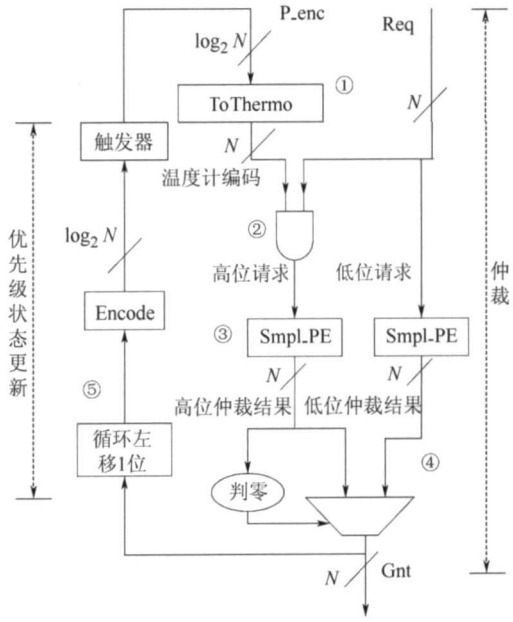

图 2 PPE的体系结构

在 PPE中, ToThermo部件实现了温度计编码算法.该算法将比特序列中指定位置两侧的比特分别置为 1和 0.例如:对于长度为 8的比特序列, 4对应的温度计编码为 11110000. 2个 Smpl_PE部件实现了固定优先级的仲裁,其中优先级自最低位向高位方向递减.

PPE的工作过程包括仲裁和优先级状态更新 2个环节.前者根据请求信号 Req和优先级状态 P_enc产生仲裁结果,后者根据仲裁结果产生新的优先级状态,实现优先级的轮转. PPE的工作过程包括 5个步骤: ①温度计编码; ②分割请求信号; ③仲裁; ④选择仲裁结果; ⑤更新优先级状态.

2 2 PPE性能分析

PPE的输出包括仲裁结果和优先级状态.在有冲突虚通道路由器和无冲突虚通道路由器中,产生上述 2个输出的关键路径分别成为 PPE的关键路径,从而体现出不同的性能.

在有冲突虚通道路由器中,位于输出端口和输入端口的仲裁器交替工作(即响应和接受步骤),每次仲裁占用一个时钟周期.此时, PPE的

关键路径是仲裁结果的产生路径,即仲裁环节涉及的前 4个工作步骤. 更新优先级状态可以在等待对方仲裁的周期中完成.

在无冲突虚通道路由器中,位于输出端口的仲裁器进行一次仲裁即可完成调度,每轮调度只占用一个时钟周期. 下一周期进行的新一轮调度需要使用本轮调度产生的优先级状态,因此仲裁器必须在同一周期内完成仲裁环节和优先级状态更新环节. 此时, PPE的关键路径是产生优先级状态的路径,包括上述所有 5个工作步骤.

综上所述,虚通道路由器的体系结构和调度机制直接影响了 PPE的性能. 在无冲突虚通道路由器中, PPE顺序产生仲裁结果和优先级状态,导致性能较低;在有冲突虚通道路由器中,由于 PPE中的仲裁环节设计不够合理,仍然存在改进空间.

3　PPE仲裁器的改进

在 PPE的基础上,本节给出 PPE_Conflict和 PPE_NonConflict 2种改进的仲裁器,分别适用于有冲突虚通道路由器和无冲突虚通道路由器.

3.1　PPE_Conflict仲裁器

在有冲突虚通道路由器中, PPE的关键路径是其中仲裁环节涉及的前 4个工作步骤. PPE存储最高优先级请求的位置作为优先级状态,在仲裁环节对其进行温度计编码. 该设计有效地减少了所需的触发器资源,但是也造成了性能的损失. 在每个仲裁周期开始后,就绪的请求信号无法立刻得到处理,而必须等待温度计编码就绪.

PPE_Conflict将 PPE中的温度计编码步骤从仲裁环节转移到优先级更新环节中,存储温度计编码作为优先级状态(图 3). PPE_Conflict在仲裁环节中直接利用保存的优先级状态进行请求信号的分割,消除了上述等待时间,缩短了产生仲裁结果的路径. 但是其所需的触发器数目从原来的 $\log_2 N$ 增加到了 N.

为了削弱上述副作用, PPE_Conflict进一步调整了仲裁环节的工作步骤. PPE利用 2个同样的 Smpl_PE部件分别对高位请求和低位请求实施仲裁,然后进行仲裁结果的选择. PPE_Conflict在仲裁前确定高位请求或低位请求作为有效请求 Valid_Req,从而只需一个 Smpl_PE部件即可产生最终的仲裁结果. 由于 Smpl_PE产生的仲裁结果非零当且仅当输入的请求非零,因此确定有效请求与 PPE中选择仲裁结果的方式效果等价,并且实现方法相同.

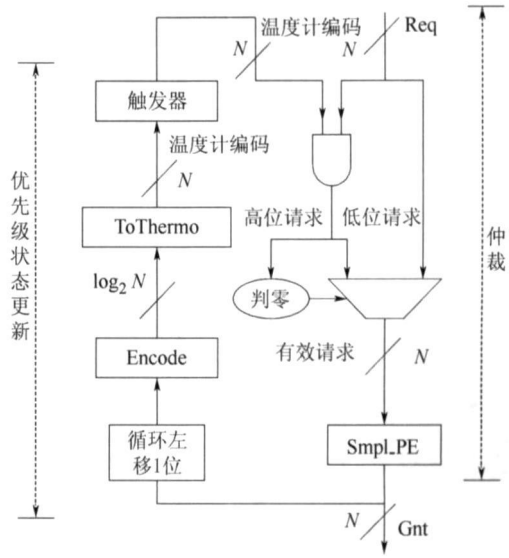

图 3　PPE_Conflict体系结构

PPE_Conflict的工作步骤包括:分割请求信号,确定有效请求,仲裁,仲裁结果编码和温度计编码. 其中前 3个步骤属于仲裁环节,后 2个步骤属于优先级状态更新环节. PPE_Conflict中的仲裁环节和优先级更新环节比 PPE更加平衡,作为关键路径的仲裁环节得到明显简化,更加符合有冲突虚通道路由器的调度特点.

3.2　PPE_NonConflict仲裁器

在无冲突虚通道路由器中, PPE的关键路径包括其中所有 5个工作步骤. PPE串行产生仲裁结果和优先级状态,明确地分割了仲裁环节和优先级状态更新环节. 该设计符合迭代式调度机制的特点,但是在无冲突路由器和非迭代调度机制中的性能较低.

PPE_NonConflict将产生仲裁结果和优先级状态的 2个步骤并行化,根据请求而非仲裁结果直接产生优先级状态的温度计编码(图 4). 其中

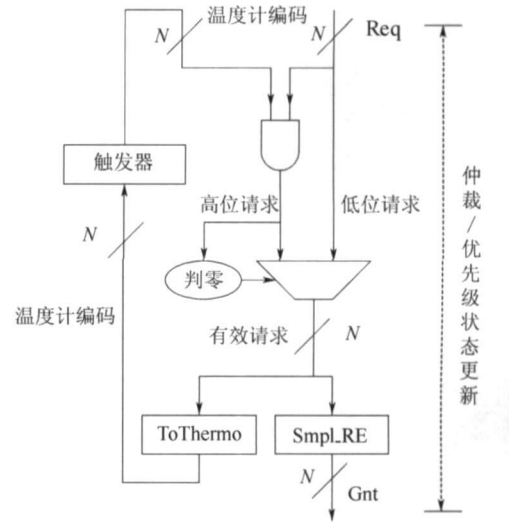

图 4　PPE_NonConflict体系结构

的温度计编码器 ToThermo实现了 PPE中 Smpl_PE, Encode和 ToThermo的全部功能. 例如: 输入请求 0101将被直接编码为 1110 与 PPE_Conflict相同, PPE_NonConflict将选择仲裁结果改为选择有效请求, 达到降低资源成本的目的.

PPE_NonConflict中的工作步骤和关键路径包括以下 3个工作步骤: 分割请求信号, 确定有效请求, 产生仲裁结果和优先级状态.

4 VLSI实现及结果分析

本文用 Verilog HDL设计了不同规模的仲裁器, 利用 Design Compiler在 TSMC 0.35μm CMOS工艺条件下进行综合. 实验结果来源于性能最优条件下的综合报告.

本节首先对 32输入的不同仲裁器的实现结果进行比较, 并分析了其中各个主要组成部分的影响; 然后进一步给出了各种规模的仲裁器的比较结果. 其中, 本文在有冲突虚通道路由器背景下对 PPE和 PPE_Conflict进行比较, 在无冲突虚通道路由器背景下进一步分析了 PPE_NonConflict的特点. 为了便于表述, PPE_Conflict和 PPE_NonConflict在本节的图表中分别表示为 PPE_1和 PPE_2

4.1 32输入的仲裁器实现结果及分析

表 1~表 4给出了 32输入的仲裁器在 2种路由器条件下的性能 (关键路径)和面积.

表 1 有冲突虚通道路由器中的面积 μm^2

仲裁器	ToThermo	Smpl_PE	Encode	触发器	总计
PPE	11200	28420	3045	1500	73988
PPE_1	4095	15155	4148	9300	59813

表 2 有冲突虚通道路由器中的性能 ns

仲裁器	ToThermo	Smpl_PE	Encode	总计
PPE	0.65	1.24	1.77	4.65
PPE_1	2.15	1.27	1.29	3.85

表 3 无冲突虚通道路由器中的面积 μm^2

仲裁器	ToThermo	Smpl_PE	Encode	触发器	总计
PPE	9135	25025	10220	1500	73603
PPE_1	12250	19055	10553	9300	78870
PPE_2	16310	17063		9300	79133

表 4 无冲突虚通道路由器中的性能 ns

仲裁器	ToThermo	Smpl_PE	Encode	总计
PPE	0.99	0.85	0.56	5.30
PPE_1	0.85	0.98	0.57	5.35
PPE_2	1.01	1.07		3.90

在有冲突虚通道路由器中, PPE中的温度计编码步骤占关键路径的 14%. PPE_Conflict通过将其移出关键路径使最终性能提高了 17%. PPE_Conflict通过调整处理流程减少了一个 Smpl_PE部件, 温度计编码器也由于仲裁环节的时间约束较弱而占用更小的面积. 因此 PPE_Conflict不但能够消除触发器数目增加带来的负面影响, 而且其总面积还优于 PPE约 20%.

在无冲突虚通道路由器中, PPE与 PPE_Conflict中关键路径的构成相同, 只是处理顺序存在差异, 因此二者具有相近的性能. 此外仅减少一个 Smpl_PE部件无法抵消触发器增多所导致的面积增加, 因此 PPE_Conflict占用的面积比 PPE高出 7%. 在上述 2种仲裁器体系结构中, 仲裁、编码和温度计编码等步骤占关键路径的 20%以上. PPE_NonConflict中的温度计编码器直接根据有效输入产生表示优先级状态的温度计编码, 并且该步骤与仲裁步骤同时进行. 结果表明, PPE_NonConflict的性能优于 PPE和 PPE_Conflict约 26.4%, 占用的面积比 PPE高出 8%, 与 PPE_Conflict的面积相近.

4.2 各种规模的仲裁器实现结果比较

图 5和图 6分别比较了不同规模的仲裁器在有冲突虚通道路由器和无冲突虚通道路由器中的实现结果.

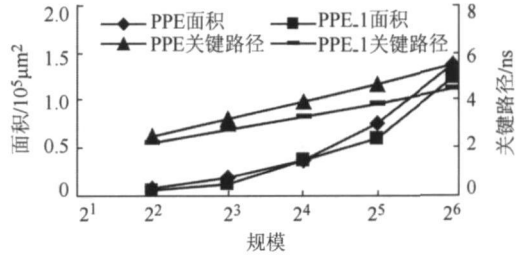

图 5 有冲突虚通道路由器中的比较

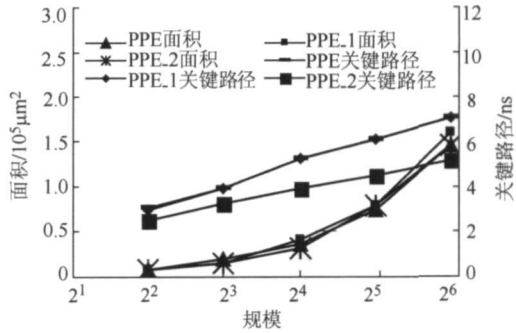

图 6 无冲突虚通道路由器中的比较

在有冲突虚通道路由器中, PPE_Conflict的性能优于 PPE约 12%~18%, 3种规模的 PPE_Conflict能够减小实现面积 15%以上 (4, 8, 32输入),

其余 2种减小 10%以下; 在无冲突虚通道路由器中, PPE_Conflict与 PPE性能相似, PPE_NonConflict的性能优于二者约 17% ~ 27%, 4种规模的 PPE_NonConflict能够减小实现面积 10%左右, 1种增加了 8% (32输入).

5 结 束 语

本文分析了虚通道路由器体系结构对调度机制和仲裁器的影响,以 PPE为基础提出了 2种改进型 PPE仲裁器——PPE_Conflict和 PPE_NonConflict二者分别比 PPE更加符合有冲突虚通道路由器和无冲突虚通道路由器. 实验结果表明, 2种改进型仲裁器在提高仲裁性能的同时, 有效地降低了实现成本.

参考文献 (**References**)

[1] Dally W J. Virtual channel flow control[J]. IEEE Transaction on Parallel& Distributed Systems, 1992, 3(2): 194-205

[2] Gupta P, McKeown N. Designing and implementing a fast crossbar scheduler[J]. IEEE Micro, 1999, 19(1): 20-28

[3] McKeown N. The iSLIP scheduling algorithm for input-queued switches[J]. IEEE/ACM Transaction on Networking, 1999, 7(2): 188-201

[4] Serpanos D N, Mountrouidou P, Gamvrili M. Evaluation of hardware and software schedulers for embedded switches[J]. ACM Transaction on Embedded Computing System, 2004, 3(4): 736-759

[5] Chao H J, Lam C H, Guo X. A fast arbitration scheme for terabit packet switches[C]//Proceedings of the IEEE Global Telecommunications Conference. Brazil: IEEE Press, 1999: 1236-1243

[6] Shin E S, Mooney III V J, Riley G F. Round-robin arbiter design and generation[C]//Proceedings of the International Symposium on System Synthesis. Japan: ACM Press, 2002: 243-248

[7] Wang F, Mounir H. Fast fair arbiter design in packet switches[C]//Proceedings of Workshop on High Performance Switching and Routing. HongKong: IEEE Computer Society, 2005: 472-476

(责任编辑:卢　硕)

(上接第 **738**页)

预期轨迹吻合较好.协同仿真结果充分验证了模型的合理性,为研究单轨车辆的动力学特性特别是稳定性提供了有效手段.

参考文献 (**References**)

[1] Cossalter V, Lot R. A motorcycle multi-body model for real time simulations based on the natural coordinates approach[J]. Vehicle System Dynamic, 2002, 37: 423-447

[3] Lai Hsien-Chung, Liu Jing-Sin, Lee D T. Design parameters study on the stability and perception of riding comfort of the electrical motorcycles under rider leaning[J]. Mechatronics, 2003, 13: 49-76

[3] 林蔚. 基于虚拟样机技术的摩托车行驶状态仿真[D]. 重庆: 重庆大学机械工程学院, 2004: 15-25
Lin Wei. Virtual prototype based simulation for motorcycle driving situation[D]. Chongqing: College of Mechanical Engineering, Chongqing University, 2004: 15-25(in Chinese)

[4] 郭孔辉. 汽车操纵动力学[M]. 吉林: 吉林科学技术出版社, 1991: 499-513
Guo Konghui. Dynamics of vehicle handling[M]. Jilin: Jilin Science& Technology Publishing House, 1991: 499-513(in Chinese)

[5] America Mechanical Dynamics Inc. MDI using ADAMS/Tire version 12.0[M]. Los Angeles: America Mechanical Dynamics Inc, 2002

[6] Brial F, Bortoluzzi D, Crossalter V, et al. Experimental study of motorcycle transfer functions for evaluating handling[J]. Vehicle System Dynamics, 2003, 39(1): 1-25

[7] Weir David H, Zellner John W. Lateral-directional motorcycle dynamics and rider control[R]. SAE 780304, 1978

[8] 吉林大学汽车动态模拟国家重点实验室. 汽车操纵稳定性模拟器试验大纲[R]. 2002
The Vehicle Dynamic Simulation National Laboratory of Jilin University. Test procedure for vehicle's handling and stability simulator[R]. 2002 (in Chinese)

(责任编辑:张　嵘)