

并发追踪数据流的多缓存选址算法

高建良, 李 欣, 王建新

(中南大学信息科学与工程学院, 湖南长沙 410083)

摘 要: 为了验证多核芯片的正确性, 通常需要同时观测不同芯核上的多组信号. 如何实时处理并发追踪中多组数据流已经成为多核芯片硅后功能验证所面临的关键挑战之一. 本文提出了一种基于映射的自调节缓存选址 (Map-Based Self-Regulation Location Selection, MSLS) 算法, 该算法通过优化多缓存选址, 在片上网络通信带宽限制下保证了并发追踪数据流能够实时存储, 同时降低了追踪数据流传输能耗. 实验结果表明了该方法的有效性.

关键词: 多核芯片; 硅后调试; 并发追踪; 多缓存选址; 片上网络

中图分类号: TN911.23 **文献标识码:** A **文章编号:** 0372-2112 (2014)11-2310-04

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2014.11.028

Multi-Buffer Location Selection Algorithm for Concurrent Trace Data Flows

GAO Jian-liang, LI Xin, WANG Jian-xin

(School of Information Science and Engineering, Central South University, Changsha, Hunan 410083, China)

Abstract: With the development of multi-core processors, it becomes a key problem to transmit concurrent trace data simultaneously to on-chip buffer under bandwidth constraint. To deal with the problem, we propose a Map-based Self-regulation Location Selection (MSLS) algorithm. This algorithm locates multiple trace buffers in interconnection fabrics under the bandwidth constraint, and reduces the average distance between trace sources and trace buffers. Experimental results show our algorithm can achieve high efficiency for post-silicon debug.

Key words: multi-core chip; post-silicon debug; concurrent trace; multi-buffer location selection; network-on-chip

1 引言

基于追踪的调试技术可以在芯片功能运行的同时将追踪数据保存到片上缓存中, 然后根据这些数据离线分析芯片内部状态^[1]. 当前一种面积开销较小的实时保存数据流的机制是复用片上网络 (Network-on-Chip, NoC) 传输追踪数据^[2]. 而通常由于预留给并发数据流的 NoC 带宽有限, 共享链路上的追踪数据会因带宽不足而丢失^[3]. 因此, 如何在有限的通信带宽的约束下提供满足调试需要的追踪信息, 成为提高多核并发追踪调试效率的关键问题^[4].

针对追踪调试的研究, 已有相关工作主要从追踪信号选择^[5~7]和追踪数据处理^[8,9]两个方面展开. 在追踪信号选择方面, 围绕着 H F Ko 等人提出的状态恢复率理论^[5], 研究人员针对门级恢复提出了一系列信号选择算法^[6,7]. 在追踪数据处理方面, 研究人员多采用字典压缩^[8]或有损压缩^[9]的方法, 利用了指令序列的连续性来对冗余信号进行压缩. 但是, 优选或压缩追踪数据的

方法无法从根本上突破带宽的瓶颈, 在应用中灵活性不足. 最近文献^[10]提出了一种基于追踪源分簇的多缓存追踪数据传输框架, 该框架将网络中的数据流量约束在簇内部, 减轻了全局链路传输压力. 而针对多缓存的安置, 作者采用了基于簇合并思想的贪心算法. 该算法适用于被调试芯核数较少时, 而当被调试芯核数较多时, 该算法无法将簇的数量合并至较小数量. 而且, 同一芯片常常需要多次反复调试不同的芯核组 (即对同一芯片存在多组追踪源), 当时文献^[10]没有考虑传输功耗的问题, 同时选址数也存在进一步优化的空间.

针对以上不足, 本文提出了一种基于多组源点映射的自调节缓存选址 (Map-based Self-regulation Location Selection, MSLS) 算法, 该算法采用映射机制将多组追踪源的信息投射到自定义的映射加权图中. 而在簇扩张的过程中, 该算法能动态调节节点的簇归属, 从而在减小缓存选址个数的同时平衡传输路径长度 (以此来降低传输能耗).

2 选址和能耗模型

2.1 多缓存选址模型与问题定义

在多缓存选址问题中,片上网络 NoC 拓扑 $G(V, E)$ 是一个有向图,其中 $v_i \in V$ 表示一个 NoC 中的路由器, $e_{u,v} \in E$ 表示节点 u 到节点 v 的链路。

定义 1 定义集合 $C\{c_i\}$ 是一组需要被调试的芯核 (即追踪源), $\varphi(c_i)$ 的值为追踪源 c_i 要发送的追踪数据量. 由具体调试参数决定。

定义 2 定义集合 $T\{t_i\}$ 是一组接收数据的追踪缓存,且每一个 $t_i \in T$ 对应一个节点簇 $Clust_i$, $Clust_i$ 至少包含一个追踪源点 c 。

定义 3 定义数据流 $flow_{u,v}$. 对 $u, v \in V$, 若有向链路 $e_{u,v} \in E$, 则 $flow_{u,v}$ 表示追踪数据从源点集 $C\{c_i\}$ 传输至终点集 $T\{t_i\}$ 时, 有向链路 $e_{u,v}$ 中的最大数据流量。

根据以上定义,多核并发调试中多缓存安置问题可以表示如下:

给定 NoC 拓扑 $G(V, E)$, 多组并发源点 $S\{C_i\}$, $i = 1, 2, \dots, n$, 链路带宽阈值 Th , 路由算法 f 。

目标 求 $\min(|T|)$ 以及 $T\{t_i\}$ 对应的 $Clust_i$, 并满足: $\forall u, v \in V, flow_{u,v} \leq Th$. $|T|$ 表示 $T\{t_i\}$ 中的缓存个数. 该问题可规约至集合覆盖问题 (NP-hard)^[11], 本文将在第 3 节详述

2.2 能耗计算模型

芯片的能耗与可靠性密切相关. 本文的能耗计算模型参考了文献[12]: 定义单比特的追踪数据从追踪源点 v 传输至追踪缓存 t 的能耗为:

$$E_{bit}^{v,t} = (hop + 1) \times (E_{in} + E_{out}) + hop \times E_{link} \quad (1)$$

其中 hop 表示追踪源点 v 至缓存放置节点 t 经过的路由器数. E_{in} , E_{out} 和 E_{link} 分别表示单比特数据在出入路由器, 以及一跳链路上所消耗的能量. 而整个追踪调试过程的总能量消耗为:

$$E_{all} = \sum_{i=1}^n \sum_{j=1}^{|V|} (E_{bit}^{v,t} \times \varphi(v_{i,j})) \quad (2)$$

其中, n 为需要独立观测的追踪源的组数, $|V|$ 表示 NoC 中路器的个数, $\varphi(v_{i,j})$ 表示第 i 组追踪源的第 j 个节点要发送的追踪数据量. 公式(1)中 E_{in} , E_{out} 和 E_{link} 是常数, 所以能耗的大小与 hop 成正比. 因此, 可以将追踪源点到追踪缓存的跳数 hop 作为数据传输的能耗指标。

3 基于映射的多缓存选址及节点成簇

3.1 多源点组的映射

对于一款实际的多核芯片而言, 通常需要进行多次调试, 每次需要观测不同的被调试芯核, 因而存在多组追踪源. 为了对相互独立的多组追踪源进行多缓存

的统一选址, 我们采用映射的方法, 由多组源点 $S\{C_i\}$ 和有向图 $G(V, E)$ 构造一个用于分簇算法的映射加权图 (Map Weighted Graph, MWG)。

定义映射加权图 $MWG(R, L, D)$, 其中 R 为 NoC 路由器节点的集合, L 为 NoC 中链路的集合, D 为 NoC 中路由器节点加权值集合. 我们假设有 n 组追踪源, 并定义 $v_{i,j}$ 为第 i ($1 \leq i \leq n$) 组追踪源映射到 NoC 中的第 j 个路由器. 进一步, d_j ($d_j \in D$) 为第 j 个路由器即节点 j 的权值, 计算公式为

$$d_j = \sum_{i=1}^n \frac{\varphi(v_{i,j}) + M(v_{i,j})}{\sum_{k=1}^{|V|} \varphi(v_{i,k})}, (j = 1 \cdots |V|) \quad (3)$$

其中 $\varphi(v_{i,j})$ 表示第 i 组追踪源的第 j 个节点要发送的追踪数据量. 函数 $M(v_{i,j})$ 为链路上数据流密集度, 表示如下:

$$M(v_{i,j}) = \frac{1}{l} \times \left(\sum_{u \in \text{neig}(v_{i,j}, f)} \varphi(u) \right) \quad (4)$$

其中 l 是节点 $v_{i,j}$ 的入度 (即相邻的双向链路的条数). 函数 $\text{neig}(v, f)$ 表示节点 v 以路由算法 f 得到的路由上一跳节点集合。

3.2 缓存选址及节点成簇

在得到图 $MWG(R, L, D)$ 后, MSLS 算法在该图上进行节点的分簇和缓存的定位. 首先我们定义集合 $N\{v\}$ 为已分簇节点, 对于 $v \in N$, 函数 $\text{head}(v)$ 计算节点 v 所归属的簇头节点 (即缓存放置节点). 在初始阶段, MWG 图中的集合 R 为全体未分簇节点. 定义函数 $\text{dist}(v, u)$ 为节点 v 和节点 u 的距离, 该距离根据路由算法 f 和 MWG 图得到。

缓存选址及节点成簇主要由三个部分组成。

(1) 簇头定位, 即缓存选址. 在开始阶段, 或者已存在的簇都因为链路负载饱和而无法继续扩张时, 从集合 D 中选择 d_v 值最大的节点 $v \in R$, 并以节点 v 作为簇头构造新的簇, 然后将该节点存入 $N\{v\}$ 中。

(2) 簇扩张. 对于一个簇 $clust_i$, 根据路由算法 f 与 MWG 图计算出簇边界节点的上一跳节点组 $Q\{v\}$. 对 $v \in Q$, 若 $v \in N$, 进入第(3)步. 否则, 若 v 加入 $clust_i$ 后, $clust_i$ 内部链路负载依然没有饱和, v 则被加入 $clust_i$ 中. 否则跳过。

(3) 簇平衡. 对 $v \in Q$, 若 $v \in N$, 取当前扩张簇的簇头 v_{new} , 并由函数 $\text{head}(v)$ 得到 v 已经分配的簇的头节点 v_{old} , 然后比较 $\text{dist}(v, v_{new})$ 和 $\text{dist}(v, v_{old})$ 的大小. 最后将 v 分配入距簇头较近的簇. 这样, 簇的大小得到了平衡。

算法的三个子过程循环进行, 一直进行到所有的簇都扩张完毕为止。

4 实验

4.1 实验设置

本部分对MSLS算法进行模拟.在实验中,NoC选取 8×8 Mesh结构和维序路由方式.在能耗计算中,本文采用文献[13]中描述的65nm TSMC的NoC架构来进行能耗的计算.具体参数如表1所示.

表1 能耗计算参数设定

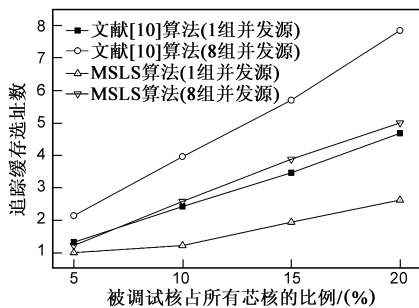
参数	设定值
出端口能耗(nW/Mb/s)	204
入端口能耗(nW/Mb/s)	94
链路能耗(nW/Mb/s/nm)	89
路由器间链路长度(mm)	2.5
路由器工作频率(MHz)	333

在实验中链路阈值 Th 设置为20位宽,观测时间为1s.为了考察在不同数据流量下算法的选址数和传输能耗.我们设计了两组实验:A组,数据流量1至 $Th/2$ 随机,即任意两个追踪源可以共享一条空闲链路.B组,数据流量 $Th/2$ 至 Th 随机,即一条空闲链路被一个数据流独占.

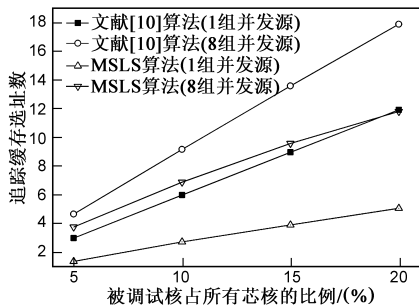
为了考察算法在一次和多次调试情况下的表现.我们分别选取在1组和8组追踪源下进行实验.被调试核的数目 k 由公式(5)确定:

$$k = \lfloor |V| \times p \rfloor \quad (5)$$

其中, $\lfloor \cdot \rfloor$ 为向下取整函数, $|V|$ 为路由器的总数量,在本文中即核的数量; p 为调试核所占比例,本文实验中 p 依次为5%、10%、15%、20%.实验数据结果取运行100



(a) 实验A



(b) 实验B

图1 选址结果比较

次的平均值.

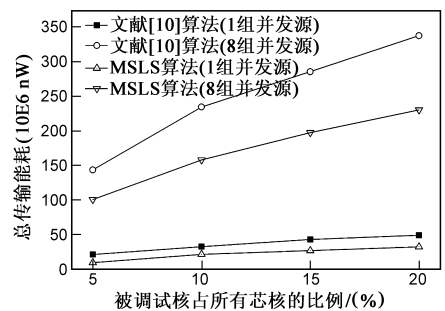
4.2 实验结果与分析

4.2.1 选址结果

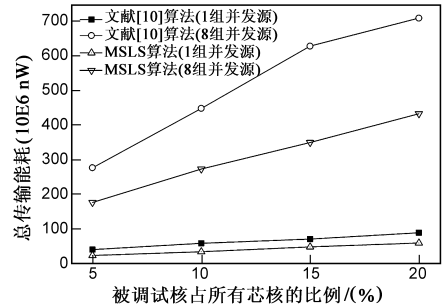
由图1可以看到,与文献[10]的算法相比,在不考虑多组追踪源变换的情况下(即1组并发源),在数据量相对较小的实验A中,MSLS算法的选址数目可以减少24%~49%;而在数据量相对较大的实验B中,MSLS算法的选址数目可以减少54%~57%.当考虑了多组追踪源变换的情况下(8组追踪源),MSLS算法的选址结果分别降低了31%~42%和20%~34%.可见实验结果与算法的设计目标一致.在实验涉及的所有情况下,MSLS都可以达到至少24%的选址数优化.

4.2.2 能耗结果

根据选址结果测定各待测核与接收缓存之间的距离,用公式(1)、(2)来计算能量消耗.图2给出了传输能耗的结果比较.在实验涉及的各种情况下,MSLS都可以比之前的算法减小30~50%的传输能耗.



(a) 实验A



(b) 实验B

图2 传输能耗比较

5 总结

本文提出了一种多缓存选址MSLS算法,通过映射来实现多组并发源条件下节点的优化,并提出自调节的簇扩张策略来实现选址节点及其簇覆盖范围的平衡.通过本文提出的算法,减少了追踪缓存选址数量,同时降低了传输能耗.实验的结果验证了该方法的有效性.

参考文献

- [1] Keshava J, Hakim N, Prudvi C. Post-silicon validation challenges: how EDA and academia can help[A]. Proceedings of the 47th Design Automation Conference[C]. Anaheim, CA: IEEE/ACM Press, 2010. 3 – 7.
- [2] Tang S, Xu Q. A multi-core debug platform for NoC-based systems[A]. Proceedings of Design, Automation and Test in Europe[C]. Nice, France: IEEE/ACM Press, 2007. 870 – 875.
- [3] Xu Q, Liu X. On signal tracing in post-silicon validation[A]. Proceedings of the 5th Asia and South Pacific Design Automation Conference[C]. Taipei: IEEE Press, 2010. 262 – 267.
- [4] Hopkins A B T, McDonald-Maier K D. Debug support for complex systems on-chip: a review[J]. IEE Proceedings-Computers and Digital Techniques, 2006, 153(4): 197 – 207.
- [5] Ko H F, Nicolici N. Automated trace signals identification and state restoration for improving observability in post-silicon validation[A]. Proceedings of Design, Automation and Test in Europe[C]. Germany: IEEE/ACM Press, 2008. 1298 – 1303.
- [6] Liu X, Xu Q. On signal selection for visibility enhancement in trace-based post-silicon validation[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2012, 31(8): 1263 – 1274.
- [7] Basu K, Mishra P. Efficient trace signal selection for post silicon validation and debug[A]. Proceedings of the 24th International Conference on VLSI Design[C]. India: IEEE Press, 2011. 352 – 357.
- [8] Lai C H, Yang F C, Huang J. A trace-capable instruction cache for cost-efficient real-time program trace compression in SoC[J]. IEEE Transactions on Computers, 2011, 60(12): 1665 – 1677.
- [9] Yuan F, Liu X, Xu Q. X-tracer: a reconfigurable X-tolerant trace compressor for silicon debug[A]. Proceedings of the 49th Design Automation Conference[C]. America: IEEE/ACM Press, 2012. 555 – 560.
- [10] Gao J, Wang J, Han Y, Zhang L, X. Li. A clustering-based scheme for concurrent trace in debugging NoC-based multi-core systems[A]. Proceedings of Design, Automation and Test in Europe[C]. Germany: IEEE/ACM, 2012. 27 – 32.
- [11] Parnas M, Ron D. Approximating the minimum vertex cover in sublinear time and a connection to distributed algorithms[J]. Theoretical Computer Science, 2007, 381(1): 183 – 196.
- [12] 杨盛光, 李丽, 高明伦, 等. 面向能耗和延时的 NoC 映射方法[J]. 电子学报, 2008, 36(5): 937 – 942.
Yang S G, Li L, Gao M L, et al. An energy and delay aware mapping method of NoC[J]. Acta Electronica Sinica, 2008, 36(5): 937 – 942. (in Chinese)
- [13] Leary G, Srinivasan K, Mehta K, et al. Design of network-on-chip architectures with a genetic algorithm-based technique[J]. IEEE Transactions on Very Large Scale Integration Systems, 2009, 17(5): 674 – 687.

作者简介



高建良 男, 1979 年出生于湖南省, 博士, 副教授, 硕士生导师. 主要研究方向为计算机系系统结构、大规模数据处理等.

E-mail: gjlpaper@gmail.com

李欣 男, 1988 年出生于安徽省蚌埠市, 硕士. 主要研究方向为多核芯片调试技术.

E-mail: leexin47@163.com

王建新(通信作者) 男, 1969 年出生于湖南省, 博士, 教授, 博士生导师. 主要研究方向为网络优化理论, 参数算法等.

E-mail: jxwang@mail.csu.edu.cn