

嵌入式处理器中分支目标缓冲器的研究与设计

王晨旭^{1,2}, 张凯峰¹, 张祥建², 喻明艳²

(1 哈尔滨工业大学(威海) 微电子中心, 山东 威海 264209; 2 哈尔滨工业大学 微电子中心, 黑龙江 哈尔滨 150001)

摘 要: 针对嵌入式应用的特点,设计了一种基于 RAM 比较 TAG 的分支目标缓冲器(BTB),并通过硬件模拟方法(BTB 控制逻辑用 RTL 实现,存储体用定制逻辑实现)研究 BTB 结构参数对 BTB 的性能、能耗以及对整个处理器系统的性能和能耗的影响,根据仿真结果选取应用于嵌入式处理器的最优 BTB 结构参数. 根据该参数,进一步设计基于 CAM 比较 TAG 的 BTB,经 SPEC2000 评测,相对于基于 RAM 比较 TAG 的 BTB,基于 CAM 比较 TAG 的 BTB 可使功耗降低 37.17%.

关键词: 分支目标缓冲器;分支预测器;嵌入式处理器;硬件模型

中图分类号: TP302

文献标识码: A

文章编号: 1000-7180(2012)01-0027-05

Research and Design of Branch Target Buffer for Embedded Processors

WANG Chen-xu^{1,2}, ZHANG Kai-feng¹, ZHANG Xiang-jian², YU Ming-yan²

(1 Microelectronics Center, Harbin Institute of Technology at Weihai, Weihai 264209, China;

2 Microelectronics Center, Harbin Institute of Technology, Harbin 150001, China)

Abstract: A RAM-tag BTB structure specifically for the application in embedded system is proposed in this paper, and the approach of hardware simulation (the control logic parts of CPU and BTB are implemented in RTL-models, while the BTB memory is simulated in circuit-model form) is employed to study the influence on the performance and the power dissipation of BTB and the whole processor imposed by BTB, finally, based on the simulation, the optimum parametric solution for the BTB structure is attained. Moreover, a CAM-tag BTB is designed based on these parameters. According to the SPEC2000 evaluation, the power consumption of CAM-tag BTB is reduced by 37.17% comparing with that of the RAM-tag BTB.

Key words: branch target buffer (BTB); branch predictor; embedded processor; hardware model

1 引言

为了提高指令运行的效率,现代处理器均采用流水结构,但在程序中的分支指令会改变指令流的方向,从而降低流水线的吞吐率,并且这种影响会随着流水线级数增加而增大,当流水线级数为 7~8 级时,吞吐率的下降幅度超过了 30%^[1],因此,分支预测技术在现代深流水处理器的性能提升中起着尤其重要的作用^[2].

现代处理器通常在分支预测器(Branch Predictor, BP)设置一个较大的分支目标缓冲器(Branch Target Buffer, BTB)来推测跳转的目标地址^[3]. 然

而在面积和能耗等因素都受限制的嵌入式处理器中,设计者往往需要在性能和实现代价等方面仔细权衡,在性能满足应用需求的前提下选择面积和能耗利用效率最高的结构参数.

本文通过硬件模拟方法(CPU 和 BTB 控制逻辑采用 RTL 模型, BTB 存储体采用电路模型)研究了 BTB 存储体的大小和相联度对处理器性能和能耗的影响,从而选取应用于嵌入式处理器的最优 BTB 结构参数. 能耗分析基于 RTL 和门级模型,相对于其它 BTB 的相关研究,结果具有更高的精确度与说服力.

BTB 存储体有两种不同的实现方式:(1)基于

RAM 比较 TAG, 它的优点是实现简单, 可配置性好; (2) 基于 CAM 比较 TAG, 它的优点是访问延迟较小, 动态功耗较低^[4]. 因此, 本文先设计了基于 RAM 比较 TAG 的 BTB 来研究 BTB 的性能和能耗以及 BTB 结构参数对处理器性能和能耗的影响, 然后根据选取的结构参数, 进一步设计基于 CAM 比较 TAG 的 BTB, 大大降低了 BTB 的动态功耗, 从而使处理器的总能耗进一步降低.

2 应用于嵌入式处理器的分支预测技术

在宽发射处理器中, 分支预测器 (BP) 一般包括方向预测器 (Direction Predictor, DP) 和 BTB 两部分, 其中 DP 具有独立的历史记录表, 并可采用 Bi-modal、两级自适应和 gshare 等多种代表性算法^[5]来提高分支预测的精度. 然而在嵌入式处理器中, 由于其发射宽度和流水级深度相对较小, 分支预测失败的代价比传统的超标量处理器要低很多, 设计者可以用一定的预测精度损失来换取与面积功耗方面的平衡. BTB 通过缓存的分支指令的相关信息来预测跳转的目标地址, 从而减小分支损失, 文献^[3]详细研究了 BTB 中存储的信息类型与执行分支指令时所需的额外时钟周期 C_{bp} 之间的关系: 当在 BTB 中同时存储分支地址 (Branch Address, BA)、分支目标地址 (Target Address, TA)、目标地址处的指令 (Target Instruction, TI) 和分支历史 (Branch History, BH) 等信息时, 可以将 C_{bp} 减小到零, 实现零分支延迟.

基于成本、芯片面积和功耗等因素的考虑, 单发射处理器仍然是嵌入式系统设计的首选, 为此, 本文的基础处理器采用单发射 7 级流水的 32 位嵌入式处理器, 其 DP 采用简单的 Bi-modal 算法, 同时将 DP 的分支历史记录表存储在 BTB 的表项中. 图 1 为基础处理器中 BP 的结构图, 它由 DP 和 BTB 两部分组成.

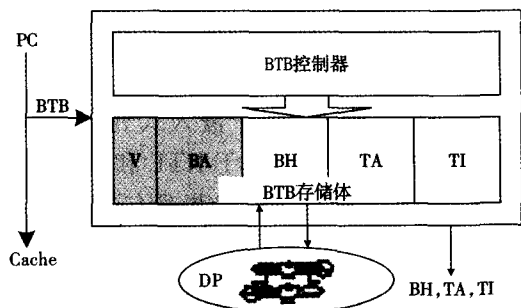


图 1 基础处理器中 BP 的结构图

3 BTB 设计与结构参数选择

3.1 BTB 模块划分及具体实现

为了确定单发射深流水嵌入式处理器中 BTB 的最优结构参数, 本文设计一个基于 RAM 比较 TAG 的 BTB RTL 模型, 并采用图 1 的 BP 结构将 DP 和 BTB 融合为统一的模块, 通过运行基准测试程序评测性能和功耗数据. 图 2 为基于 RAM 比较 TAG 的 BTB 结构框图, 主要由两部分组成: 控制器和 BTB 存储体, 其中, 控制器负责 BTB 的检索、更新、填充和替换控制.

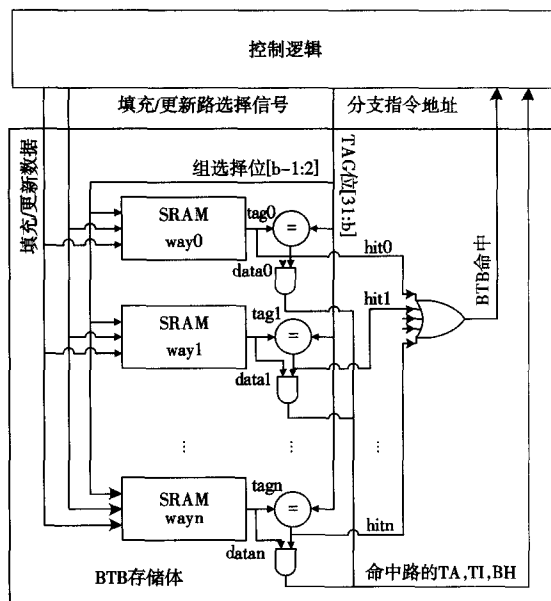


图 2 基于 RAM 比较 TAG 的 BTB 结构框图

BTB 设计可分为以下几个功能模块:

(1) 检索控制: 处理器的 PC 低位 $[b-1:2]$ 为 BTB 组地址, 它选择 n 路组相联 BTB 存储体中 n 个 SRAM 中的分支项, PC 的高位 $[31:b]$ 为比较 TAG, 它用于比较所有组相联 SRAM 输出的分支项中 BA 域是否有相等的值, 若相等则表示分支命中, 输出命中信号 hit 和命中的路 n , 以及该分支项的数据 TA、TI 和 BH 域.

(2) 更新控制: 当分支项命中后, BTB 将命中路 n 和分支项中的信息缓存, 当命中的分支指令运行到执行级后, 根据执行的结果与分支项中的 BH $[1:0]$ 来更新分支历史信息, BH 的更新状态机如图 3 所示. 之后将更新后的信息根据命中路 n 和 BA 的 $[b-1:2]$ 位存入 BTB 存储体的 SRAM 中.

(3) 填充与替换控制: 对于不发生跳转的分支, 即使不做任何预测也不会带来性能损失, 因此, 本设

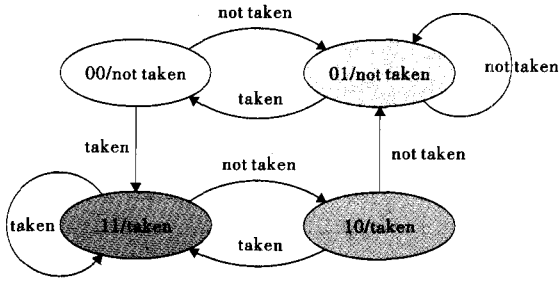


图3 基于Bimodal算法的BH更新状态机

计使用的填充策略是分支指令在BTB中没有记录并发生跳转时才触发填充BTB的动作,填充时将有效位V置位,分支历史域设为初始状态弱跳10.若在BTB填充时所有项的有效位V都已置位,则需要替换操作,常用的替换算法有随机法、先入先出(FIFO)、最不经使用算法(LFU)、最近最少使用算法(LRU)和轮循法(Round Robin).对于小型的高相联度BTB,前几种替换算法的实现代价都较高,因此本文采用实现较简单的Round Robin替换算法.在n路组相联的BTB中,每组的所有分支项共享一个 $\log_2(n)$ bit的轮循指针,指针指向该组要替换的路,替换完成后指针加一指向下一个要替换的路.

(4) BTB存储体:在n路组相联的BTB中,BTB存储体由n个SRAM的阵列构成,本文仿真和功耗分析的SRAM模型均由SMIC 0.18 μ m工艺的Memory compiler生成.SRAM中的分支项包括:(1)比较TAG,包括有效位V和分支地址BA;(2)数据,包括分支目标地址TA、目标地址处的指令TI和分支历史BH.各项数据的位宽为:V为1bit;BA的低位[b-1:2]为组地址,因此只存储其高位[31:b];TA只存储30bit字地址;TI为32bit;由于DP采用两位饱和计数器,只需2bit来存储历史信息,另外用1bit高位来标识条件分支与无条件跳转指令,因此BH为3bit.

3.2 BTB结构参数选择

为了获取在嵌入式处理器中性能与功耗较优的BTB结构,本文通过改变BTB的相联度(BTB ways)和组数(BTB sets)得到不同的BTB大小(BTB size),分别对直接映射、2路、8路和32路组相联的32、128、512和2048分支项的BTB进行了实验.

3.2.1 仿真平台

本文采用的基础处理器包含一个性能监测模块,实验通过性能监测模块获得检索和命中等数据以计算BTB性能;同时从RTL仿真的波形文件

(VCD, FSDB)提取中开关活动性文件(SAIF),然后用Power Compiler工具分析BTB与处理器的动态功耗,功耗分析模型采用SMIC 0.18 μ m工艺库的门级网表与CAM和SRAM电路模型.

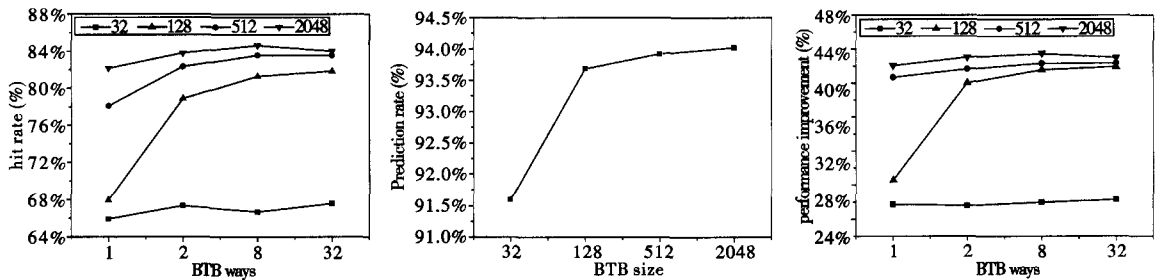
3.2.2 测试程序信息

在采用RTL模型仿真获取原始数据过程中,运行整个基准程序所需的时间的代价很大,一般都采用运行部分代码的方式来加速仿真实验,较通用的做法是采用X+Y方式的截断执行(truncated execution)进行快速模拟,文本中,X=20 M,Y=100 M,即预热20 M指令后详细模拟100 M指令.同时,我们保证分支特性研究的完备性的前提下减小了基准程序集,即根据分支指令数、基本块大小、跳转的分支指令数和后向跳转数等分支特性和程序的相似性选取了SPEC2000中的七个程序(crafty、equake、vortex、gcc、art、vpr-place和gzip等)作为测试集.

3.2.3 实验数据分析

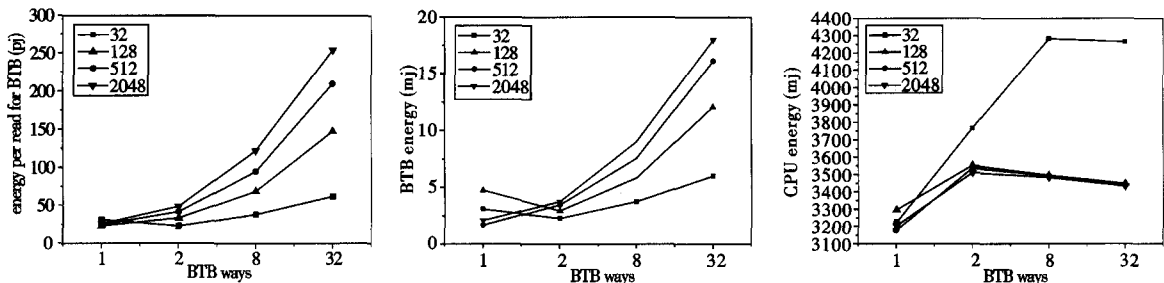
当BTB大小固定时,分支项被替换的概率随相联度的增加而减小,因此命中率上升;相联度越大,被替换的概率减小趋势变缓,导致命中率上升的幅度变缓;另外,BTB增大时,命中率也将增大,增大的趋势同样也逐渐变缓,如图4(a)所示.DP的历史信息存储在BTB的分支项中,当BTB增大时,分支项被替换的概率减小,BH的更新状态机能维持稳定状态,因而DP预测精度也随之提高;当预测精度增加到一定程度时,分支项替换的影响变小,因此预测精度上升趋势变缓,如图4(b)所示.受BTB命中率和DP预测精度的影响,处理器性能(IPC)的提升也随BTB大小和相联度增加而增大,变化趋势与BTB命中率类似,如图4(c)所示;可以看出,128分支项BTB对性能提升的影响随相联度变化最明显,同时,32路组相联时的性能接近于512和2048分支项的BTB,因此,32路组相联128分支项的BTB以较小的实现代价就可以达到较好的性能.

由于TAG比较需要用到大量的比较逻辑,随着相联度的增加,BTB每次检索时消耗在TAG比较上的能量迅速增大;同时,相联度固定时,每次检索时访问SRAM所消耗的能量也随BTB的增大而增加,如图5(a)所示.同理,BTB消耗的总能量也有类似的变化趋势;当相联度和BTB大小增加时,总能耗也随之而增大,如图5(b)所示.但另一方面,分支预测对于处理器性能的提升也随着BTB相联度和大小的增加而增大,因此,从图5(c)中可以看到,



(a) BTB 的命中率随大小和相联度的变化 (b) DP 预测精度随分支项大小的变化 (c) BP 对处理器性能提升的影响随 BTB 大小和相联度的变化

图 4 性能数据



(a) BTB 每次检索消耗的能量随大小和相联度的变化 (b) BTB 消耗的总能量随大小和相联度的变化 (c) 处理器消耗的总能量随大小和相联度的变化

图 5 能耗数据

除了直接映射的 BTB, 在运行本实验的测试程序时, 处理器的总能耗随 BTB 的相联度和大小的增加而减小. 在直接映射 BTB 中, 由于不需要进行 TAG 比较, 处理器的总能耗也相对较小. 拥有 32 分支项 BTB 的处理器总能耗受 TAG 比较能耗的影响较大, 所以总能耗随相联度的增加而增大.

综合图 4 和图 5, 可以看出, 由于 BTB 对处理器性能具有提升作用, 它将使处理器运行固定负载时花费的时间更少, 因此, 即使在 BTB 上多花费一些能耗也是值得的, 它的存在最终将使处理器总能耗更低. 根据实验数据, 32 路组相联 128 分支项的 BTB 对于性能提升有较好的表现, 此外, 它可使处理器的面积和总能耗达到较低的水平. 综合考虑嵌入式处理器的应用需求以及 BTB 存储体的实现代价和对处理器性能、能耗的影响, 32 路组相联 128 分支项的 BTB 是单发射深流水嵌入式处理器的最优配置.

4 基于 CAM 比较 TAG 的 BTB 设计

根据 BTB 组相联的特性, BTB 命中率将随相联度的增加而增大, 此外, 进行 TAG 比较时的延迟也会随之增大, 以致增加其动态功耗, 而将 BTB 的存储体用基于 CAM 比较 TAG 的存储结构来实现

可以缩短 TAG 的比较时间, 从而大大降低 BTB 的功耗. 因此, 根据前面确定的 BTB 结构参数, 本文还设计了基于 CAM 比较 TAG 的 BTB 存储体, 如图 6 所示.

存储体主要由三部分组成: (1) 用 CAM 实现的比较 TAG; (2) 编码解码器; (3) 用 RAM 实现的数据存储阵列. 当 BTB 填充或更新时, PC 的低位 $[a-1:2]$ (组地址) 选择 BTB 存储体某个组 setn, 填充/更新路选择信号经编解码器后选择 RAM 中相应的路, 然后将要填充或更新的数据写入存储体中; 当 BTB 检索时, 若输入的 PC $[31:a]$ 与某路比较 TAG 的所有位都相等, 匹配线 ml 将被拉高, 经编解码器编码后选中 RAM 中的对应分支项数据 (TA, TI 和 BH), 然后将其放大输出. 各路 ml 相或后输出命中 hit 信号, 同时所有 ml 信号编码后输出命中的路 n.

高相联度的 BTB 可以通过基于 CAM 比较 TAG 的存储结构来有效降低动态功耗^[6], 经前面所述测试程序评测, 32 路组相联 128 分支项 BTB 的平均动态功耗为 11.16mw; 相比较而言, 相同大小的基于 RAM 比较 TAG 的 BTB 平均动态功耗为 17.76mw. 也就是说, 基于 CAM 比较 TAG 的 BTB 使功耗降低了 37.17%, 如图 7 所示.

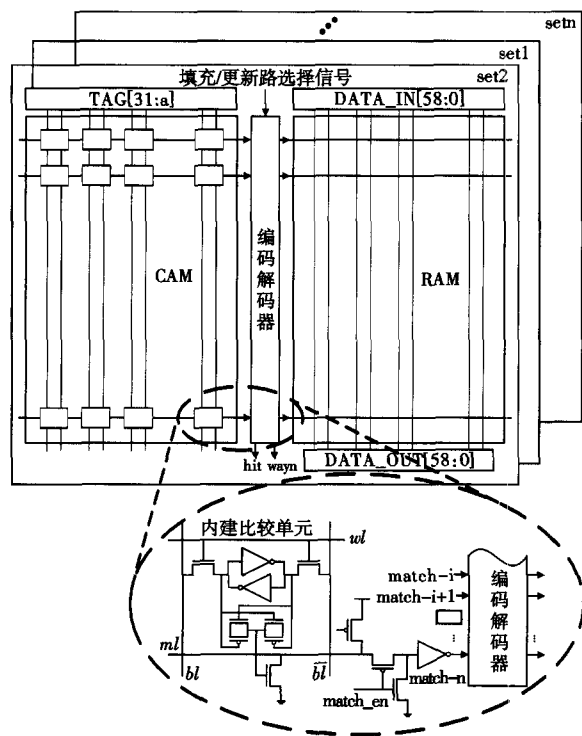


图 6 基于 CAM 比较 TAG 的 BTB 存储结构框图

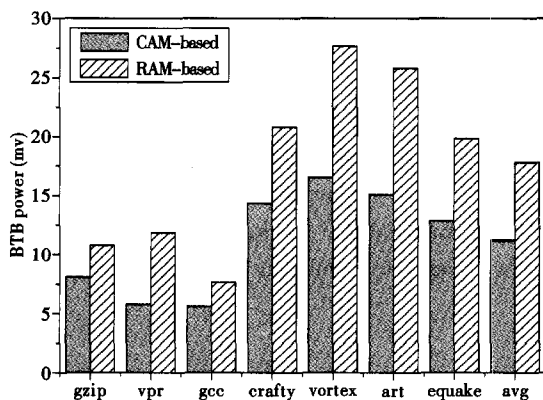


图 7 基于 CAM 与基于 RAM 比较 TAG 的 BTB 功耗

5 结束语

本文设计了一种基于 RAM 比较 TAG 的 BTB,并通过硬件模拟的实验方法选取了应用于嵌

入式处理器的最优 BTB 结构参数. 为了降低 BTB 存储体的动态功耗, 本文进一步设计了基于 CAM 比较 TAG 的 BTB, 经 SPEC2000 评测, 相对于基于 RAM 比较 TAG 的 BTB, 基于 CAM 比较 TAG 的 BTB 可使功耗降低 37.17%.

参考文献:

- [1] 郑纬民. 计算机系统结构 [M]. 2 版. 北京:清华大学出版社, 1998: 512-514.
- [2] Evers M, Yeh TY. Understanding branches and designing branch predictors for high-performance microprocessors [J]. Proceedings of the IEEE, 2001, 89 (11): 1610-1620.
- [3] Perleberg C, Smith A. Branch target buffer design and optimization [J]. IEEE Transactions on Computers, 1993, 42(4): 396-412.
- [4] 喻明艳, 张祥建, 王晨旭. 分支指令特性与分支预测器的性能研究[J]. 微电子学与计算机, 2010, 27(6): 8-12.
- [5] Scott McFarling. Combining branch predictors[R]. California: Digital Western Research Laboratory, 1993.
- [6] 鄢国丰, 张盛兵, 黄小平. “龙腾 R”微处理器分支处理单元的研究与设计[J]. 微电子学与计算机, 2010, 27 (6): 122-127.

作者简介:

王晨旭 男,(1977—),博士研究生.研究方向为嵌入式系统、数字系统设计.

张凯峰 男,(1986—),硕士研究生.研究方向为数字系统设计.

张祥建 男, (1984—), 硕士研究生. 研究方向为处理器体系结构与 SoC 设计.

喻明艳 男,(1965—),教授.研究方向为处理器体系结构与 SoC 设计.