

# A Reconfigurable Cache Memory with Heterogeneous Banks

Benitez, D. ; Moure, J.C. ; Rexachs, D. ; Luque, E.; DATE 2010

**Aluno:** John Edward Esquiagola Aranda RA: 124208

O tamanho ótimo da memória cache num SoC (*System on chip*) pode ser diferente dependendo dos programas a serem executados nele. Num mesmo ponto, a redução dos erros de cache conseguidos quando aumentarmos o tamanho da memória, tem como consequência um aumento no tempo de latência de cache, afetando o desempenho da mesma. A melhor configuração de cache depende das características da aplicação e das restrições do projeto (desempenho, potência e área). Devido ao fato de que não existem caches que satisfaçam os requerimentos de todas as aplicações, uma abordagem promissora é adicionar capacidade de reconfiguração na memória cache.

Este artigo em questão apresenta o *Amorphous Cache (AC)*, uma memória cache L2 reconfigurável desenvolvida para melhorar o desempenho assim como reduzir o consumo de potência. AC é composta de sub-caches heterogêneas que podem ser desligadas seletivamente, reduzido o tamanho total da cache e/ou sua associatividade, o qual reduz o tempo de acesso e o consumo de potência. Uma técnica de reconfiguração de cache chamada *BBV&NDP* também é proposta, essa técnica usa *Basic Block Vectors (BBVs)* para identificar as fases de execução de um programa dinamicamente. *Amorphous Cache*, é um circuito reconfigurável especializado que pode ser implementado na tecnologia CMOS e pode ser integrado nos processadores convencionais. O termo *Amorphous* indica que o range de configurações da cache, caminhos críticos e consumo de potência não é homogêneo, eles dependem da carga de trabalho.

A arquitetura consiste em 18 configurações de AC, indo desde 64KB até 2MB, e com associatividade em grupos de 4, 8 ou 16 vias. O circuito AC é organizado num conjunto de blocos heterogêneos chamados *Sub-Caches*, os quais são conectados nas entradas e saídas da cache através de uma inovadora rede de interconexão programável dentro do chip. As sub-caches que não estiverem conectadas são desligadas. Existem registradores de configuração que permitem modificar o tamanho e a associatividade do AC, Quando uma configuração pequena é requerida, só a sub-cache com o tamanho requerido é ativado e usado, e o restante é desligado usando 1-bit num registrador de configuração de *Power-Supply*.

Quando um código de programa é executado, ele vai por diferentes estágios. O processo de reconfiguração começa limpando a *sub-cache* activa ou todas as *sub-caches*. Só depois, a sub-cache requerida ou todas as *sub-caches* são activadas. Tendo como alvo detectar dinamicamente as mudanças nos estágios do programa e executar cada estágio com a máxima eficiência se propõe um novo algoritmo de hardware called *BBV&NDP*. Este algoritmo resume a execução do programa coletando informação sobre quais e quantas vezes os blocos básicos de instrução (*BB*) são executados durante um intervalo de instrução, a partir deste ponto um novo vector chamado '*BBV reduzido*' é criado a partir dos 300 *BBs* mais executados.

Esses vectores *BBV* reduzidos são armazenados na *Pattern CAM Memory*, eles são chamados de *Pattern BBVs* e representam os diferentes estágios do programa. Usa-se o algoritmo *NDP* para determinar qual *Pattern BBV* será designado para a configuração da cache.

De acordo com os autores, a heterogeneidade do *Amorphous Cache* faz dela uma proposta diferente se compararmos com o estado da arte atual dos caches reconfiguráveis. O novo predictor baseado em *BBV* proposto selecciona a melhor configuração de cache se comparado com outras técnicas adaptativas. Os resultados demonstram que o AC reduz a latência de acesso, potência dinâmica e potência estática na

media de 55,8%, 46,5% e 49,3% quando comparado com uma memória cache não adaptativa de 2MB.