

A Cache Memory System based on a Dynamic/Adaptive Replacement Approach

José Aguilar *

Ernst Leiss[†]

Resumen

En este trabajo, nosotros proponemos un sistema de memoria cache basado en un esquema de reemplazo adaptativo, el cual formaría parte del Sistema Manejador de la Memoria Virtual de un Sistema Operativo. Nosotros usamos un simulador de eventos discretos para comparar nuestro enfoque con trabajos previos. Nuestro esquema de reemplazo adaptativo esta basado en varias propiedades del sistema y de las aplicaciones, para estimar/escoger la mejor política de reemplazo. Nosotros definimos un valor de prioridad de reemplazo a cada bloque de la memoria cache, según el conjunto de propiedades del sistema y de las aplicaciones, para seleccionar cual bloque eliminar. El objetivo es proveer un uso efectivo de la memoria cache y un buen rendimiento para las aplicaciones.

Palabras Claves: *Sistema de Manejo de Memoria, Memoria Cache, Evaluación de Rendimiento.*

Abstract

In this work we propose a cache memory system based on an adaptive cache replacement scheme, as part of the virtual memory system of an operating system. We use a sequential discrete-event simulator of a distributed system to compare our approach with previous work. Our adaptive cache replacement scheme is based on several criteria about the system and applications with the objective being to estimate/choose the best replacement policy. We assign a replacement priority value to each cache block according to a set of criteria to select which block to remove. The goal is to provide an effective utilization of the cache memory and good application performance.

Keywords: *Memory Management System, Cache Memory, Performance Evaluation.*

1. Introduction

In high-performance computer systems, memory bandwidth is often a bottleneck. There are two ways to optimize the memory access, the transformation of the code to adapt to the cache/memory system or the definition of optimal memory management systems (for example, a dynamic replacement policy). Cache memory is the simplest cost-effective way to achieve a high speed memory hierarchy. A cache provides, with high probability, instructions and data needed by the CPU at a rate that is more in line with the CPU's demand rate. Three basic cache organizations have been defined at the level of cache memory [1, 2, 4]: direct mapped, fully associative and set

* CEMISID. Dpto. de Computación, Facultad de Ingeniería. Universidad de los Andes. Merida - Venezuela 5101. aguilar@ula.ve

[†] Department of Computer Science. University of Houston. Houston, TX 77204-3475, USA. cosce@cs.uh.edu