CS-221 Computer Architecture Project

A Performance Study of Software And Hardware Data Prefetching Mechanisms

---

Collaborators :

RITIK JAIN 18114068
RAHUL SAHANI 18114062
PRATEEK SACHAN 18114056
ANKITESHWAR D ATRI 18116011
VIKAS UPADHYAY 18116083
MEHUL ARYA 18116048

Abstract

Prefetching, i.e., exploiting the overlap of processor computations with data accesses, is one of several approaches for tolerating memory latencies. Prefetching can be either **hardware-based** or **software-directed** or a combination of both.

Hardware-based prefetching, requiring some support unit connected to the cache, can **dynamically handle** prefetches at run-time without compiler intervention.

Software-directed approaches **rely on compiler technology** to insert explicit prefetch instructions.

In this paper, we discuss the two representative schemes : software-based prefetching and hardware-based prefetching, and evaluate approximations to these two schemes in the context of a **shared-memory multiprocessor environment**.
We qualitatively compare, of both schemes, the ability to reduce **cache misses** in the **domain of linear array references.**

When complex data access patterns are considered, the software approach has **compile-time information** to perform sophisticated prefetching whereas the hardware scheme has the **advantage of manipulating dynamic information.**

The performance results from an instruction level simulation of some standard benchmarks are included for comparison.

An approach combining software and hardware schemes shall be proposed; which promises memory latency reduction with least overhead.