

A Prefetching Memory System for Mediaprocessors

Stefan G. Berg¹, Donglok Kim² and Yongmin Kim³

¹Department of Computer Science and Engineering

²SVision LLC, Bellevue, WA

³Departments of Electrical Engineering and Bioengineering

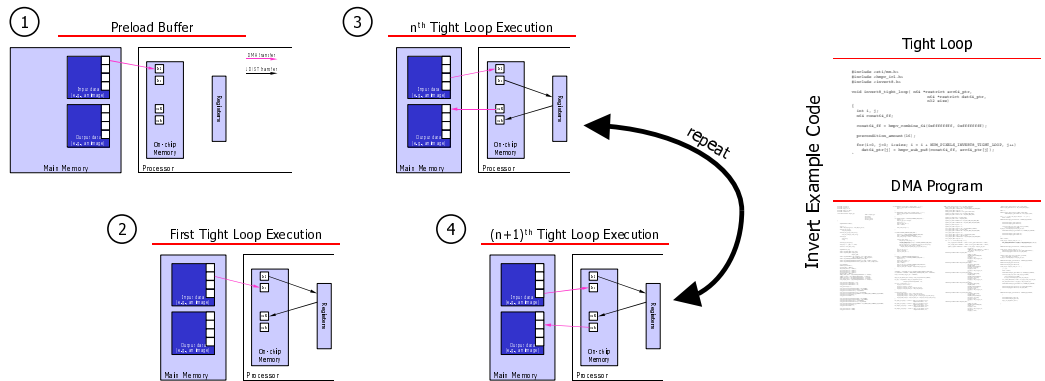
Mediaprocessor Overview

- Programmable processor
 - Flexible
 - Cost-effective (~ US\$50)
 - High-performance
- Multimedia processing
 - Set-top boxes, DVD players, printers, camcorders, HDTV, security devices, camera on a chip, etc.
- Examples: Hitachi/Equator Technology MAP-CA, Texas Instruments TMS320C64x, TriMedia Technologies TM-64
- Likely to be a dominant choice for future embedded market

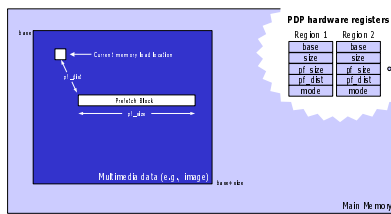
Mediaprocessor Programming Compared

- | | |
|---|--|
| <p>Cache-based</p> <ul style="list-style-type: none"> - Performance: <ul style="list-style-type: none"> + Register pressure and data dependencies limit ability to overlap computation and data transfers - Little control to efficiently utilize main memory - Programmability: <ul style="list-style-type: none"> + Cache is transparent to programmer, no DMA program needed | <p>DMA-based</p> <ul style="list-style-type: none"> - Performance: <ul style="list-style-type: none"> + Computation and data transfer effectively overlapped + Blocked memory transfers typically effectively utilize DRAM's page-mode access, thus achieving high throughput - Programmability: <ul style="list-style-type: none"> - DMA programs can be difficult to write and debug - Lack of standard API for DMA programming complicates porting of functions to other architectures |
|---|--|

Double Buffering with Direct Memory Access (DMA)



Program-directed Prefetching (PDP)

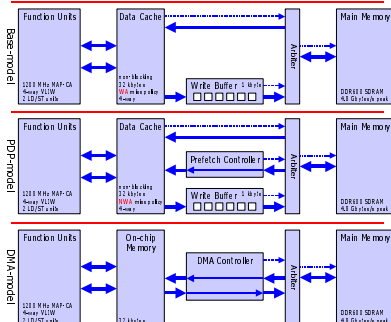


- Allows simple cache-based programming
 - Programmer must only specify a set of prefetch hints per region of data, which are stored in hardware registers
- Mimics data flow of DMA controller to achieve DMA-like performance
 - Prefetching of larger blocks of data possible to efficiently access main memory
 - Can prefetch far ahead for effective overlapping of computation and memory transfer
 - Multiple prefetch regions used to maximize prefetch accuracy for each type of data flow present in function

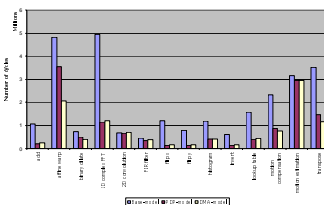
No-write-allocate (NWA) Write-miss Cache Policy

- Write-allocate (WA) policy popular, but not efficient for writing output data
- NWA eliminates allocation of output data in cache when output data is not modified
- For typical image processing functions with one input image and one output image, WA has 50% greater memory traffic than NWA

Simulation Models



Execution Time Results



Conclusions

- DMA-based programming: best performance
- Cache-based programming: fast development
- Program-directed prefetcher:
 - Runs cache-based programs annotated with prefetching hints
 - Achieves almost DMA-based performance
 - Less flexible than DMA controller
 - Can simplify porting of functions across platforms
- PDP is intended to significantly simplify software development process without sacrificing performance