

文章编号 :1001 - 2486(2004)02 - 0046 - 05

基于硬件计数器的 MPEG4 视频应用性能特征分析*

车永刚¹, 王正华¹, 李晓梅²

(1. 国防科技大学并行与分布处理国家重点实验室, 湖南 长沙 410073; 2. 装备技术指挥学院, 北京 101416)

摘要 通过指令集扩展来增强多媒体应用性能是微处理器发展的一个趋势, Intel MMX、SSE 和 SSE2 是很好的例子。但应用本身的特点是否支持这些体系结构特性, 应用软件是否充分发挥了多媒体指令集的潜力, 都是值得关注的问题。采用动态 Profiling 的方法, 研究 MPEG4 视频应用在基于 Intel P4 处理器的桌面机上的性能特性, 使用硬件计数器工具来获取指令级的性能数据, 并进行了深入分析。发现 MPEG4 播放软件较好地利用了 Intel P4 的 64 位 MMX 指令集, 但对 SSE 和 SSE2 指令集几乎没有利用上, 两级 Cache 不命中导致的性能损失很大。为提高视频播放性能, 改进存储层次设计与优化播放软件比改进多媒体指令集更加有效。

关键词 MPEG4 视频; 硬件性能监视; 性能特征; MMX; SSE; SSE2

中图分类号: TP302.7 文献标识码: A

Hardware Counter-based Performance Characterizing for MPEG4 Video Applications

CHE Yong-gang¹, WANG Zheng-hua¹, LI Xiao-mei²

(1. National Lab. for Parallel and Distributed Processing, National Univ. of Defense Technology, Changsha 410073, China;

2. Institute of Equipment and Command Technology, Beijing 101416, China)

Abstract It is the current trend that microprocessors extend their instruction set to speedup the multimedia applications. Intel's MMX, SSE and SSE2 are good example. But do the multimedia applications favor these architecture features? And does the current software support these instruction sets and make full use of them? These questions are of much concern to the researchers. We investigate the performance characteristics of 2 MPEG4 Video players on an Intel P4 based desktop PC by means of dynamic profiling. Based on Intel Vtune, a hardware performance monitoring tool, we obtain accurate performance data. And we make in depth analysis of the obtained data based on the machine's architecture parameters and performance profiles. We find that the 2 applications make good use of 64-bit MMX instructions. But they rarely use SSE and SSE2 instructions. The performance degradation due to Cache misses at tow levels of Caches is remarkable. To enhance the performance of MPEG4 media applications, improving memory hierarchy performance and optimizing the media players are more effective ways than employing more advanced multimedia instruction sets.

Key words MPEG4 videos; hardware performance monitoring; performance characteristics; MMX; SSE; SSE2

视频播放需要处理大数据流, 并进行大量解码运算, 对处理器性能提出了很高的要求。为满足视频处理等多媒体应用的需求, 处理器厂商也对其处理器体系结构进行了改进, 如 Intel 就扩展了其处理器指令集, 先后推出了 MMX、SSE、SSE2 技术。然而, 这些技术的实际效果如何? 视频播放软件能否从应用特点与机器体系结构特征出发, 充分发挥机器的运算处理潜力? 这些都是人们非常关心的问题。就目前流行的 MPEG4 视频应用来说, 人们对其播放器软件与实际平台间的相互作用仍然缺乏深入研究, 特别是缺乏定量分析。多数研究工作采用基准程序测试方法, 考察的是系统级的综合性能指标, 关注对器件性能的对比评价, 而没有从体系结构与应用特征之间的关系入手进行分析。

对 Intel P4 处理器 + Microsoft Windows 平台上两种新兴 MPEG4 视频应用的性能特征进行研究, 目的是考察 MMX、SSE 和 SSE2 在 MPEG4 视频应用中的效果, 探究体系结构与应用特征之间的匹配程度, 定位性能瓶颈, 以期对体系结构发展和播放软件设计有所启发。与基于静态模拟或基准程序测试的方法不

* 收稿日期: 2003 - 09 - 01

基金项目: 国家自然科学基金重点资助项目(69933030)

作者简介: 车永刚(1973—), 男, 博士生。

同, 使用动态性能特征分析(profiling)方法, 针对程序的真实执行过程进行研究。

1 方法

1.1 视频播放器

使用两种流行的视频播放器: Microsoft Windows Media Player 9.0 (简称 Wmp) Real Networks Real One Player 2.0 (简称 Real)。另外使用了 Free Software Foundation 的 Xvid MPEG4 Video 解码器、VobSub 2.23 和 AC3Filter 0.70b 等辅助软件。

1.2 视频媒体文件

使用 DivX 编码的 avi 文件、Microsoft 的 wmv 文件, 它们预先存储在本机硬盘上, 基本数据如表 1。

表 1 视频文件

Tab.1 The video media files

媒体格式	压缩标准	视频比特率	帧/s	分辨率	播放器
Divx (.avi)	MPEG4	1.5Mbps	25	578 × 326	Wmp, Real
Wmv (.wmv)	MPEG4 改进型	1495.2Kbps	24	636 × 476	Wmp, Real

在下文中, wmv_Divx 表示使用 Wmp 播放 Divx 媒体的情况, 其他类似。

1.3 测试平台

1.3.1 系统配置与性能指标

表 2 给出了实验平台的软硬件配置与性能指标。操作系统是 Windows 2000 Professional。

表 2 实验平台的硬件配置与性能指标

Tab.2 The hardware configuration and performance metrics of the experiment platform

CPU	Intel P4 2.4GHz	主板	Intel D865Perl (800M FSB)
L1 Cache	12KB trace /8KB data /64B/line, 理论带宽 76.8GB/s	硬盘	Maxtor 80GB ATA133
L2 Cache	512KB /128B/line, 理论带宽 76.8GB/s	显示卡	GeForce4 MX440 - 8x
主存	512MB 266MHz DDR SDRAM, 理论带宽 2128 MB/s	声卡	AC97

1.3.2 MMX、SSE与SSE2^[1]

MMX (MultiMedia eXtension)指令是 64 位打包的单指令多数据 (Single Instruction Multiple Data, SIMD) 整数指令, 以并行方式处理 8、16 或者 32 字节的多个数据元素, 从而提高软件运行速度。SSE (Streaming SIMD Extension)引入了 70 条新的加速浮点计算和 3D 计算的指令, 这些指令使用专门的 128 位寄存器。SSE2 在 SSE 基础上加入了一系列新指令, 能够对打包的多个数据项在单个寄存器中进行操作, 也可以同时对多个数据进行操作。

Intel P4 处理器中兼具 MMX、SSE 与 SSE2 指令集。应用软件要利用这些指令集, 需要使用支持这些指令集的编译器 (如 Intel C/C++ 编译器) 或特殊的函数库。

1.4 基于硬件性能监视机制的性能数据获取

硬件性能监视 (Hardware Performance Monitoring) 利用处理器中内置的硬件性能计数器 (Hardware Performance Counters) 来监视与统计处理器内与性能密切相关的事件的发生, 获得程序运行期间的性能数据 (如周期与指令计数、存储层次访问数量及命中情况等), 为体系结构部件的能效分析与程序性能优化提供指导。基于硬件计数器的性能测试获得的性能数据精确可靠^[2], 而且系统开销小^[3]。目前大多数高性能处理器中都设置了硬件性能监视硬件。

Intel P4 中有 18 个性能计数器^[4]。通过 Intel Vtune Performance Analyser 7.0^[5]来访问它们。Vtune 基于事件的采样 (Event-based Sampling, EBS) 提供了利用硬件计数器进行程序性能分析的功能。

1.5 数据测试

使用 Vtune 的 EBS 会话 (activity) 选择需要测试的播放程序 (以媒体文件的全路径及其文件名为其命令行参数), 设置其测试周期为 20s, 并设置等待 10s 后再开始进行数据采集, 每个性能事件的采样周期

结束后停止播放器。然后,选择希望测试的性能事件,设置对来自操作系统与用户级的性能事件都进行计数。启动该会话后,Vtune对每个性能事件分别执行下面的步骤:

- (1) 启动播放器,播放器打开相应的媒体,按照默认的速度和窗口大小进行播放;
- (2) 等待一段设定的时间(10s),然后开始进行事件采样;
- (3) 测试周期到,停止播放器,将测试得到的数据写入缓冲区。

由于软/硬件的限制,Vtune在一次采样期间实际上只能监控1种或少数几种性能事件(多数情况下是1种)。设置保证Vtune测得的不同性能事件都是该播放器在相同的时间内播放相同媒体文件的相同部分期间产生的。同时,为避免干扰,在测试时确保系统中无其他活跃的应用程序。

2 实验结果及分析

2.1 指令比例

各种指令所占的比例是程序性能特征的重要指标。表3列出了影响性能的主要指令所占的比例。

表3 各种指令所占百分比

Tab.3 The instruction components(%)

	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
Load	27.19	30.23	44.76	27.78
Store	15.81	18.16	21.59	13.39
64-bit MMX	29.22	15.44	10.39	24.76
Branch	6.30	9.53	4.71	7.75
Float Point	5.56	6.78	1.81	2.25
128-bit MMX	0.0092	0.0131	0.0043	3.2930
Packed Single-precision FP SSE	0.0019	0.0028	0.1933	0.0018
Packed Double-precision FP SSE(SSE2)	0	0	0.0103	0

可以看到,除了Wmp_wmv,其他情况都没有使用SSE2指令,而Wmp_wmv中SSE2指令比例也只有0.0103%。对4种情况,SSE指令所占的比例都非常低,只使用了打包的单精度SSE指令,而且比例都在0.2%以下。除了在Real_Wmv中达3.293%以外,128位MMX指令所占比例也非常少。这说明应用几乎没有使用SSE2指令,对128位MMX指令和SSE指令的利用也非常有限。而64位MMX指令所占比例较大,说明应用很好地利用了64位MMX指令。浮点指令数也占较大比例,说明应用中的很多浮点操作并没有被向量化。此外,Load与Store指令的比例最高,说明应用的寄存器重用很少,多数运算指令所需的操作数都需要通过访存来获得。

2.2 指令执行速率

指令执行速度是重要的性能指标。表4只列出3个参数——IPC(Instruction Per Cycle):指令每周期;64b MMX/Cyc:64位MMX指令每周期;FP/Cyc:浮点指令每周期。在计算这些值时,只计算花在播放进程上的周期数,因为实验中播放器的CPU占用率都小于100%,故该时间只是采样周期的一部分。

表 4 指令执行速率

Tab.4 The instruction execution rates

	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
IPC	0.298	0.247	0.667	0.304
64b MMX/Cyc	0.087	0.038	0.069	0.059
FP/Cyc	0.017	0.017	0.012	0.011

Ins/Cyc 都小于 0.7, 相对于 CPU 最大流出率 3 指令/每周期来说小很多, 指令执行的效率很低。64b MMX/Cyc 的值都在 0.05 ~ 0.21, FP/Cyc 的值在 0.011 ~ 0.017, 都大大低于处理器的峰值性能。

2.3 Cache 不命中率

由于 CPU 速度与存储器速度的差距日益扩大, Cache 命中率对程序性能影响非常显著。表 5 给出了 L1 Data Cache Load 的不命中率、L2 Cache Load 的不命中率、Trace Cache 不命中率。

表 5 Cache 不命中率(%)

Tab.5 Cache miss rates(%)

	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
L1 Cache Load Miss Rate	9.40	10.12	3.21	8.00
L2 Cache Load Miss Rate	0.88	0.76	0.23	0.44
Trace Cache Miss Rate	0.31	0.38	0.06	0.20

除了 Wmp_wmv 之外, L1 数据 Cache 加载的不命中率都大于 5%, 说明程序在 L1 Cache 的访存局部性并不令人满意。而 L2 Cache 加载的不命中率非常小, 可能因为该级 Cache 大小已经能够容纳应用的工作集。Trace Cache 的不命中率非常低, 说明应用程序有很好的指令局部性。

2.4 存储带宽

现代计算机中存储带宽与 CPU 处理能力不匹配的问题也很突出, 所以存储带宽需求也是程序性能分析必须关注的问题。根据 Vtune 参数与机器体系结构参数, 计算得到各级存储带宽需求(表 6)。

表 6 各级存储带宽需求(单位: MB/s)

Tab.6 Memory bandwidth requirements for 3 memory hierarchies (MB/s)

	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
L1 data Cache	2459	2296	8491	2853
L2 Cache	1169	1161	1473	1134
Main memory	803	578	465	409

与表 2 比较, 可看出各级存储带宽需求在绝对大小都小于机器的各级存储的理论带宽。

2.5 程序平衡

文献 [6] 中提出了浮点程序平衡(Program Balance)、机器平衡(Machine Balance)的概念。如果程序平衡大于机器平衡, 则程序执行期间由于存储带宽不足, 满足不了 CPU 计算所需的数据供应, 程序性能受限于存储带宽。因为在 MPEG4 播放应用中, 64 位 MMX 指令完成主要的视频与音频解码工作, 访存也主要是为 64 位 MMX 指令提供数据, 故对其平衡概念作一定的修改。

定义 1: 执行每条 MMX 指令平均需要访问的存储字节数, 称为程序平衡。

定义 2: 机器执行每条 MMX 指令期间平均能够传输的存储字节数, 称为机器平衡。

程序平衡与机器平衡都是针对特定存储层次而言的。经过计算得到各级平衡值, 如表 7 所示。

表7 各级存储上的程序平衡与机器平衡

Tab.7 Program balances and machine balances at 3 memory hierarchies

	机器	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
L1 data Cache	32	9.89	17.42	43.51	16.99
L2 Cache	32	4.70	8.81	7.55	6.75
Main memory	0.887	3.23	4.39	2.38	2.43

对 L1 数据 Cache 除了 Wmp_wmv 的程序平衡大于机器平衡外,其他 3 种情况下程序平衡都小于机器平衡。对 L2 Cache 程序平衡都小于机器平衡。而在主存这一级,程序平衡都大于机器平衡很多。可见在多数情况下,L1 Cache 和 L2 Cache 的存储带宽都不成为 MPEG4 视频播放应用的性能瓶颈。但是在主存一级,程序平衡都大大高于机器平衡,这一方面因为主存带宽与 P4 计算能力相比差距巨大,另一方面也可能说明应用程序的访存与计算比较多。

2.6 分支预测

分支指令对程序性能有重要影响,一方面因为分支指令影响了流水线的正常执行序列,降低了 CPU 可用的指令级并行性;另一方面,分支预测错误会带来额外的开销(如重新取指、分支目标缓冲、BTB 的更新等)。特别是 P4 采用长达 20 级的流水线,分支预测错误的开销更大。从表 2 中可以看出,分支指令在全部指令中占了相当的比例。表 8 列出了分支预测错误率(Branch Miss Rate)。

表8 分支预测错误率(%)

Tab.8 Branch miss-prediction rates(%)

	Wmp_Divx	Real_Divx	Wmp_wmv	Real_wmv
Branch Misprediction Rate	6.81	6.33	6.97	6.42

分支预测错误率都在 6.33% ~ 6.97%,这一比例相对较高。

2.7 性能影响率

Vtune 除了统计事件发生次数之外,还估计了部分事件导致性能损失比例,如 L1 Cache 加载不命中中的性能影响。根据 Cache 开销模型与机器的性能配置,估计了 L2 Cache 不命中中的性能影响率。图 1 给出了几种重要的性能事件对性能的影响率。

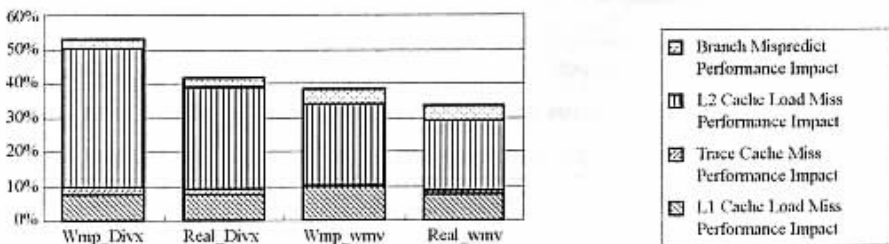


图1 各种事件的性能影响率

Fig.1 Performance impact of several performance events

L2 Cache Load Miss 的开销最大,在 20% ~ 41%,原因是主存带宽相对较低,加上 P4 的 L2 Cache 行长很长(128 字节),使得 L2 Cache 不命中开销很大。L1 Cache Load Miss 的开销在 7% ~ 10%。分支预测错误导致的性能损失也在 2.5% ~ 4.5%。Trace Cache 不命中导致的性能损失非常小。可见,对 P4 上的 MPEG4 应用,性能瓶颈主要在主存带宽、L1 Cache 不命中和分支预测上。

如在具体应用中时, BP 算法或多或少带有经验的成分; 一个成功的应用还要建立在深入了解应用背景的基础上; 不仅需要正确地选择网络模型、网络结构和学习算法, 还需要根据试验结果适当地调整众多参数, 在网络投入实际使用之前, 往往还有大量的数据选择、训练、变换、调整和维护工作需要完成。

参考文献:

- [1] 丁建立, 杜二明, 许卫明. 马尔可夫预测模型及其在市场预测中的应用[J]. 洛阳师专学报, 1997, 4(2): 42-45.
- [2] 刘丽萍, 秦婷. 回归分析在饮料销售量预测中的应用[J]. 预测天地, 2001, 6(28): 28-29.
- [3] 李吉成, 沈振康. 灰度形态滤波的神经网络实现[J]. 系统工程与电子技术, 1999, 3(56): 56-59.
- [4] 黄德双. 神经网络模式识别系统理论[M]. 北京: 电子工业出版社, 1996.
- [5] 庄镇泉, 王熙法, 王东生. 神经网络与神经计算机[M]. 北京: 科学出版社, 1991.

(上接第 50 页)

3 结论

(1) MPEG4 视频应用未能充分利用 P4 处理器的 SSE2 指令和 SSE 指令, 但较好地利用了 64 位 MMX 指令。此外, 未向量化的浮点指令仍占一定比例。所以, 视频播放软件进一步面向体系结构进行优化很有必要, 也有很大潜力。

(2) 两级 Cache 不命中导致的性能损失占很大比例。针对视频应用, 处理器设计应改善 Cache 的容量与延迟特性, 同时提高机器主存性能, 这可能比进一步增强处理器中的多媒体指令集更加经济有效。

参考文献:

- [1] Intel 公司. IA-32 Intel Architecture Software Developer's Manual Volume 1: Basic Architecture[R]. 2001.
- [2] Schaelicke L, et al. Profiling Interrupts in Modern Architectures[C]. Eighth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, Los Alamitos, CA, 2000: 115-123.
- [3] Hirzel M, Chilimbi T. Bursty Tracing: A Framework for Low-overhead Temporal Profiling[C]. Workshop on Feedback-directed and Dynamic Optimizations, 2001: 117-126.
- [4] Intel 公司. IA-32 Intel Architecture Software Developer's Manual Volume 3: System Programming Guide[R]. 2001.
- [5] [Http://developer.intel.com/vtune/](http://developer.intel.com/vtune/). 2003.
- [6] Carr S. Combining Optimization for Cache and Instruction-level Parallelism[C]. PACT '96, Boston, 1996: 238-247.

