

CA 4 存储层次作业

姓名：丁彦添 学号：1911406

1. (20) 假设如下的Cache容量/延迟时间/MPKI(misses per thousand instruction): 32KB/1/100, 128KB/2/80, 512KB/4/50, 2MB/8/40, 8MB/16/10, 访问片外内存平均需要200个时钟周期。对于下述Cache配置, 计算访问Cache层次的平均访问时间。对于太浅或者太深的Cache层次, 你观察到它们的缺点是什么?

a) 32KB L1; 8MB L2; 片外内存

$$MissPenalty_{L2} = HitTime_{Memory}$$

$$= 200 \text{ cycles}$$

$$MissPenalty_{L1} = HitTime_2 + MissRate_2 * MissPenalty_{L2}$$

$$= 16 \text{ cycles} + \frac{10}{1000} * 200 \text{ cycles} = 18 \text{ cycles}$$

$$AccessTime = HitTime_1 + MissRate_1 * MissPenalty_{L1}$$

$$= 1 \text{ cycle} + \frac{100}{1000} * 18 \text{ cycles} = 2.8 \text{ cycles}$$

对于太浅的 Cache 层次, 有了更多的可能去访问片外存储, 降低了整体延迟时间。

b) 32KB L1; 512KB L2; 8MB L3; 片外内存

$$MissPenalty_{L3} = HitTime_{Memory}$$

$$= 200 \text{ cycles}$$

$$MissPenalty_{L2} = HitTime_3 + MissRate_3 * MissPenalty_{L3}$$

$$= 16 \text{ cycles} + \frac{10}{1000} * 200 \text{ cycles} = 18 \text{ cycles}$$

$$MissPenalty_{L1} = HitTime_2 + MissRate_2 * MissPenalty_{L2}$$

$$= 4 \text{ cycles} + \frac{50}{1000} * 18 \text{ cycles} = 4.9 \text{ cycles}$$

$$AccessTime = HitTime_1 + MissRate_1 * MissPenalty_{L1}$$

$$= 1 \text{ cycle} + \frac{100}{1000} * 4.9 \text{ cycles} = 1.49 \text{ cycles}$$

c) 32KB L1; 128KB L2; 2MB L3; 8MB L4; 片外内存

$$MissPenalty_{L4} = HitTime_{Memory}$$

$$= 200 \text{ cycles}$$

$$MissPenalty_{L3} = HitTime_4 + MissRate_4 * MissPenalty_{L4}$$

$$= 16 \text{ cycles} + \frac{10}{1000} * 200 \text{ cycles} = 18 \text{ cycles}$$

$$MissPenalty_{L2} = HitTime_3 + MissRate_3 * MissPenalty_{L3}$$

$$= 8 \text{ cycles} + \frac{40}{1000} * 18 \text{ cycles} = 8.72 \text{ cycles}$$

$$MissPenalty_{L1} = HitTime_2 + MissRate_2 * MissPenalty_{L2}$$

$$= 2 \text{ cycles} + \frac{80}{1000} * 8.72 \text{ cycles} = 2.697 \text{ cycles}$$

$$AccessTime = HitTime_1 + MissRate_1 * MissPenalty_{L1}$$

$$= 1 \text{ cycle} + \frac{100}{1000} * 2.697 \text{ cycles} = 1.269 \text{ cycles}$$

对于太深的 Cache 层次, 结构设计复杂, 成本较高, 功耗增高, 但延迟提升不明显, 性价比低。

2. (20)您正在设计一台PMD并优化它以降低能耗。处理器核，包括一个8KB L1数据缓存，在不处于休眠状态时的能耗1W。如果处理器核有一个完美的L1缓存命中率，那么对于给定任务，它将达到CPI=1，也就是说，执行1000条指令需要1000个周期。每个附加访问L2和L2之外缓存的周期都会给处理器核增加一个停顿周期。基于下述说明，对于给定任务，L2缓存的容量多大才能使PMD（处理器核，L1缓存，L2缓存，内存）具有最低能耗？

- a) 处理器核频率是1GHz，L1缓存的MPKI是100。
b) 256KB L2缓存的延迟是10个周期，MPKI是20，背景功耗为0.2W，每次L2访存消耗0.2nJ（纳焦耳）。
c) 1MB L2缓存的延迟是20个周期，MPKI是10，背景功耗为0.8W，每次L2访存消耗0.7nJ。
d) 内存的平均延迟是100个周期，背景功耗0.5W，每次访问内存消耗35nJ。

对于处理器核来说，每个时钟周期的时间为 $\frac{1}{1GHz} = 1ns$ 。

L2缓存的平均额外功耗为：

L2缓存背景功耗+L1 缓存缺失率*(1+L2缓存缺失率)*L2缓存背景功耗+L1缓存缺失率*L2缓存缺失率*访问内存功耗*(1+L1缓存缺失率)

256KB L2缓存的平均额外能耗为：

$$0.2W + 10\% * \frac{1010}{1000} * \frac{0.2nJ}{1ns} + 10\% * 2\% * \frac{35nJ}{1ns} * \frac{1100}{1000} = 0.2972W$$

1MB L2缓存的平均额外能耗为：

$$0.8W + 10\% * \frac{1020}{1000} * \frac{0.7W}{1ns} + 10\% * 1\% * \frac{35nJ}{1ns} * \frac{1100}{1000} = 0.9099W$$

由计算得知：选择 256KB 的 L2 缓存能耗更低。

3. (15)在运行多个程序的处理器中，最后一级缓存通常由所有程序共享，这会导致干扰，即一个程序的行为和缓存占用会影响其它程序对缓存的使用。首先，从服务质量（QoS）的角度来看，这是一个程序，其中干扰导致程序拿到的资源比承诺的更少，性能也比承诺的更低。其次，这是一个隐私问题。根据所看到的干扰，程序可以推断其它程序的内存访问模式，这被称为定时通道，一种从一个程序泄漏到另一个程序的信息形式，可被用来危害数据隐私或反向工程竞争对手的算法。您可以向上一级缓存中添加哪些策略，以便一个程序的行为不受共享缓存的其它程序行为的影响？

- 解决性能影响：对数据存储的分配不由初始的状态决定，而是基于程序运行中反馈的实时状态来调节针对每个应用程序的缓存使用量。
- 解决安全问题：为不同程序设置关键数据存取区，保护重要的数据。

4. 您正在设计一个写透L1缓存和写回L2缓存之间的写缓冲器，L2缓存写数据总线宽度是16字节，并且每四个处理器周期可以完成对独立的缓存地址执行一次写操作。

- a) (15) 写缓冲区中每个条目的宽度应该是多少个字节？

16字节

- b) (15) 如果所有其它指令都可以与存储指令同时发出，并且所需的数据块在L2缓存中，则通过执行64位存储指令清零存储器时，使用合并写缓冲区代替非写缓冲器，在稳定状态下预期的加速比是多少？

合并写缓冲区需要 $16 \div 8 = 2$ cycles，非合并写缓冲区需要 4cycles，加速比为 $\frac{4}{2} = 2$

- c) (15) 对于拥有阻塞和非阻塞缓存的系统，L1未命中会对所需写缓冲区条目的数目产生什么影响？

对于阻塞缓存不会产生变化，但对于非阻塞缓存，由于有写缓冲区的处理机制，未命中时需要的条目会更少。

