

实验4：构建cache模拟器

1 实验目的

实验目的：

- 完成cache模拟器
- 理解cache块大小对cache性能的影响；
- 理解cache关联性对cache性能的影响；
- 理解cache总大小对cache性能的影响；
- 理解cache替换策略对cache性能的影响；
- 理解cache写回策略对cache性能的影响。

2 实验任务

2.1 任务

2.1.1 实现的功能

编写程序，实现以下功能：

- 读取配置文件，按配置文件实现cache模拟器；
- 对给出的跟踪文件进行回放，计算并输出cache模拟器的性能参数。

2.1.2 命令行的使用

编译生成cachesim程序，典型用法包括

```
cachesim -c cfg.txt -t ls.trace [-o ls.trace.out]
```

其中：

- -c 后面的参数指定配置文件，配置文件的格式见2.2；
- -t 后面的参数指定跟踪文件，跟踪文件的格式见2.3；
- -o 后面的参数指定输出文件，输入文件的格式见2.4；该参数是可选的(意味着该参数可以不写)，此时直接输出到标准输出。

2.2 配置文件

缓存模拟器将指定的参数是指定包含所有必需的配置文件的参数参数来设置缓存并计算需要输出的统计信息。配置文件的布局具有必须遵守的特定格式。它包含6行，顺序如下：

- 块大小：以字节为单位的缓存块(或cache line)大小，该值应该是2的幂次；
- 关联性：指定缓存的关联性。值“1”表示直接映射缓存，而值“0”表示完全关联，可以为2的非负幂次，代表 2^k 组相联；
- 数据大小：指定缓存中数据的总大小。这不包括任何开销（例如标记大小），以KB为单位且是2的非负幂。例如，值“64”表示64KB缓存。
- 替换政策：指定要使用的替换策略。值“0”表示随机替换，值“1”表示LRU,其余值无效；
- 非命中开销：指定缓存未命中时的额外周期数，可以是任何正整数；

- 写入分配：在缓存非命中时的策略。值“0”表示非写分配，值“1”表示写入分配，其余值无效。

下面的配置文件，指定了一个16KB的直接映射缓存，其中8字节块、LRU替换策略、100周期未命中开销以及遵循写入分配策略。

```
8
1
16
1
100
1
```

2.3 跟踪文件

跟踪文件的部分信息如下

```
0x7fd2e03921b0: R 0x7fd2e03905b4
0x7fd2e03921b5: R 0x7fd2e03905b6
0x7fd2e03921ba: R 0x7fd2e03905b8
0x7fd2e03921cd: R 0x7fd2e03be9e8
0x7fd2e0392260: R 0x7fd2e0390f90
0x7fd2e0392264: W 0x7fd2e03be038
0x7fd2e0392187: R 0x7fd2e0390fa0
0x7fd2e039218b: R 0x7fd2e0390f98
```

该跟踪文件使用pin-3.27[1]在linux下面生成。每一行的格式如下

指令地址： R/W 数据内存地址

其中：

- 指令地址：一个48位整数（采用无符号十六进制格式），指定正在访问的指令地址。例如，“0x7fd2e03921b0”指定访问内存地址140543681700272。
- 访问类型：代表对内存的访问方式是读('R')还是写('W')；
- 数据地址：一个48位整数（采用无符号十六进制格式），指定正在访问的指令地址。

2.4 模拟器输出

模拟器应将“.out”扩展名写入输出文件文件，其前缀基于输入跟踪文件。例如，如果您输入“gcc.trace”，则输出文件应命名为“gcc.trace.out”。模拟器将计算将进入输出文件的几个统计信息。就像配置文件一样，输出文件每行应包含一项数据，它应包含以下5行：

- 总命中率：命中的内存操作(即跟踪文件中的行)的百分比，其数字应保留小数点后4位；
- load命中率：命中的load百分比，其数字应保留小数点后2位；
- store命中率：命中的store百分比，其数字应保留小数点后2位；
- 总运行时间：程序的总运行时间，其单位应为周期数；
- 平均内存访问延迟：完成内存访问所需的平均周期数，其数字应保留小数点后2位；

输出的形式应该如下(请忽略后面的数值)：

Total Hit Rate: 98.15%
Load Hit Rate : 99.50%
Store Hit Rate: 97.10%
Total Run Time: 12341341234
AVG MA Latency: 150.20

3 实验提交

3.1 实验报告

实验报告的内容包括：

- 设计的cache模拟程序的说明：包括主要的数据接口，主要的函数调用（篇幅不能多余1页），以及典型的输入/输出；
- 不同cache大小情况下，不同cache块大小对命中率的影响。cache块大小可以取32、64、128、256字节，数据cache大小取16KB、32KB、64KB、128KB或256kb，直接映射、写分配、LRU替换规则；画一幅图，类似于PPT 85页的图。
- 固定cache大小情况下，不同关联性对命中率的影响。数据cache大小取128KB，关联性选择直接映像、2路组相联、4路组相联、8路组相联，写分配、LRU替换规则；画一幅图。
- (可选内容)不同替换策略对于cache的影响。选择数据cache大小为128KB，4路组相联、写回法，比较随机替换法和LRU替换法的性能差异。
- (可选内容)不同写回策略对于cache的影响。选择数据cache大小为128KB，直接映像、LRU替换法，比较写回法和写分配法的性能差异。

3.2 源代码

提交完整的源代码文件。

4 关于回写策略的说明

在实验参数中，只有一个参数来说明回写策略：值“0”表示非写入分配，值“1”表示写入分配，其余值无效。

- 实际上，在cache写命中的时候，有两种策略，分别是写回法(Write back)，即只写入cache；写直达法(Write Through)，即同时写入cache和主存。两者的差别是写回法会导致cache和主存的不一致；而写直达法不会导致cache和主存的不一致。
- 在cache写不命中的时候，也有两种策略，分别是非写分配法和写分配法。非写分配法在不命中时，直接写主存；写分配法在写不命中的时候，尝试所写的内存内容调入cache中(可能需要替换)，然后再在cache中写一次(此时为写命中)。两者的差异是非写分配中主存和cache是一致的，但是在写分配中cache和主存是不一致的。
- 命中有两种策略，不命中也有两种策略，总的组合是4种方式。但在实际中，从cache和主存一致的角度出发，只有两种方式：即写命中时的写直达和写不命中时的非写分配--这两种情况都保持了cache和主存的一致性；写命中时的写回法和写不命中时的写分配--这两种情况都保持了cache和主存的不一致性。
- 因此只需要一个参数即可。

5 参考资料

[1] pin tools 3.27 用户文档。!(pin-tools-3.27)

[<https://software.intel.com/sites/landingpage/pintool/docs/98718/Pin/doc/html/index.html>]

[2] cache的基本原理。《计算机组成原理与体系结构》第3版本，7.5高速缓冲存储器。