volatile 实现原理

volatile 关键字保证了内存的可见性,但是不保证原子性。

volatile 为访问变量加入了内存屏障,防止指令重排序。

被 volatile 修饰的变量在编译之后的指令中,定义变量的 flags 会加上 ACC_VOLATILE 标志。

ACC_VOLATILE 标志被定义在 JVM 源码的 accessFlags.hpp 中。

```
1 // accessFlags.hpp
2 bool is_volatile () const { return (_flags & JVM_ACC_VOLATILE) != 0; }
4 // bytecodeInterpreter.cpp
5 // 存储变量时, 判断是否被 volatile 修饰
6 if (cache -> is_volatile()) {
      // 判断数据类型,根据不同的数据类型执行不同的方法
     if (tos_type == itos) {
         // int 类型
         obj -> release_int_field_put(field_offset, STACK_INT(-1));
1.0
     } else if (tos_type == atos) { // obj 对象类型
          VERIFY_OOP(STACK_OBJECT(-1));
12
          obj -> release_obj_field_put(field_offset, STACK_OBJECT(-1));
13
          OrderAccess::release_store(&BYTE_MAP_BASE[(uintptr_t)obj >> CardTableModRefBS::card_shift];
      } else if (tos_type == btos) {
15
          // byte 类型
          obj -> release_byte_field_put(field_offset, STACK_INT(-1));
17
     } else if (tos_type == ltos) {
18
          // long 类型
          obj -> release_long_field_put(field_offset, STACK_LONG(-1));
20
     } // ... char, short, float, double 省略
21
      // 执行完毕后,执行下面这个方法
      OrderAccess::storeload();
23
24 }
25
26 // oop.inline.cpp
27 // release_int_field_put 方法在此文件中
28 inline void oopDesc::release_int_field_put(int offset, jint contents) { OrderAccess::release_store(
30 // release_store 方法在 orderAccess.hpp 中定义
31 // orderAccess.hpp
32 static void release_store(volatile jint* p, jint v); //还有对其他数据类型的定义
34 // 具体的实现根据不同的操作系统CPU进行实现 Linux, Windows.. 等等 CPU
35 // 例如: orderAccess_linux_x86.inline.hpp
36 inline void OrderAccess::release_store(volatile jint* p, jint v) { *p = v; } // 此处 volatile, 语言级
```

- 1. 对每个 volatile 变量的写操作的前面会插入 storestore barrier。
- 2. 对每个 volatile 变量的写操作后会插入 storeload barrier。
- 3. 对每个 volatile 读操作之前插入 loadload barrier。
- 4. 对每个 volatile 读操作之后插入 loadstore barrier。

以上代码第 23 行证实了第二点: (其他的可以在其他源码中找到)

```
5 inline void OrderAccess::storestore() { release(); }
6 inline void OrderAccess::loadstore() { acquire(); }
7 inline void OrderAccess::storeload() { fence(); }
8
9 // fence() 方法中会在指令之前加入 lock 前缀。(汇编指令)。。。汇编指令没有深入研究,不贴代码了。
```

所以 volatile 修饰符可以保证内存的可见性(内存屏障)。但是无法保证高并发下的原子性。例如:

```
package com.keanu.io.study.concurrency;

public class VolatileDemo {

volatile int i = 0;

public void incr() {
    i++;
    }

public static void main(String[] args) {
    new VolatileDemo().incr();
    }
}
```

被 JVM 编译之后(使用 javap -c 指令查看字节码):

```
public class com.keanu.io.study.concurrency.VolatileDemo {
volatile int i;
public com.keanu.io.study.concurrency.VolatileDemo();
   Code:
      0: aload_0
                           // Method java/lang/Object."<init>":()V
      1: invokespecial #1
      4: aload_0
      5: iconst_0
                                      // Field i:I
       6: putfield #2
1.0
       9: return
11
public void incr();
14 Code:
      0: aload_0
15
       1: dup
16
                             // Field i:I
       2: getfield
17
       5: iconst_1
18
       6: iadd
19
                                      // Field i:I
20
       7: putfield
                      #2
      10: return
21
22
  public static void main(java.lang.String□);
23
    Code:
24
       0: new
                                      // class com/keanu/io/study/concurrency/VolatileDemo
25
       3: dup
26
       4: invokespecial #4
                                      // Method "<init>":()V
27
       7: invokevirtual #5
                                      // Method incr:()V
28
      10: return
29
30 }
31
```

可以看到,编译之后的指令中,i++(复合操作) 的操作分成了 3步,以上代码的 17, 19, 20 行。

- 1. aetfield
- 2. iadd
- 3. putfield

当有多个线程同时执行时,有可能同一时间有多个线程同时执行了 getfield 指令,可能就会有一个线程拿到的是旧值,这就造成了原子性问题。

可以通过 synchronized 关键字来解决,避免线程并行执行。synchronized 实现原理可以参照这篇文

章: https://keanu96.github.io/deep-in-synchronized/