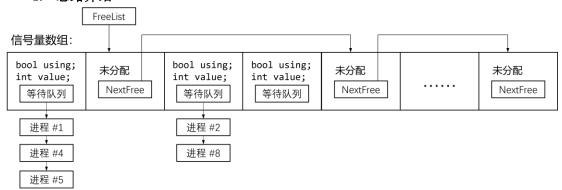
# 第3次操作系统实验报告

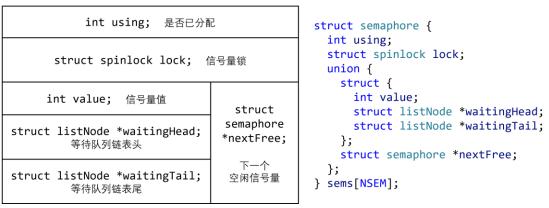
161810129 董世晨

### 一、 信号量

#### 1. 思路介绍



#### Semaphore结构体:



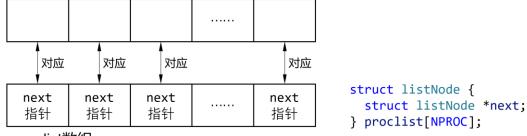
图注:纵向堆叠表示struct,横向堆叠表示union。

上图包含了我认为设计中比较关键的地方,相信别人看后应该比较容易复现我的方案。接下来我介绍一下设计中需要注意的细节。

#### 2. 如何实现等待队列

首先,我们不希望出现饥饿的情况,队列结构可以符合这个要求。另外观察到:每个进程或者不等待,或者正在等待一个信号量,不可能出现同时等待多个信号量的情况。**最简单的方式是在PCB中新增一个字段,记录等待队列列表的下一个。** 

#### ptable.proc数组:



proclist数组:

但是由于存在奇怪的要求,PCB中不能新增字段,因此我开辟了一个和ptable.proc大小相同的数组proclist,其中元素和ptable.proc中的元素按下标一一对应,如上图所示。例如,proclist[5]对应了ptable.proc中的下标为5的PCB,注意不是和PID为5的进程对应。这样子,listNode作为链表的结点,作为进程的替身,被加入到了等待队列中。

#### 3. 锁

除了每一个信号量有一个细粒度锁来保证wait和signal时不发生数据竞争,还需要一个大粒度锁来保护整个信号量数组在索引、分配和释放信号量时不发生数据竞争。另外,在修改PCB内容和进入调度器时,需要锁定ptable,并释放已获取的锁。

比起为什么需要加锁,我觉得更值得思考的是**为什么不需要加锁**。在我的设计中,不需要对proclist加锁,也就是说,**如果多个进程同时执行wait\_sem,但wait的具体信号量不同,它们能并发执行。**只有wait\_sem和wait\_signal这两个函数读取和修改对proclist中的内容,下面讨论它们并发执行的三种情况:

- 1) 两个进程同时执行wait\_sem: 由于ptable.proc和prolist中的元素——对应,这两个进程对应的proclist元素必然不同,因此不会发生冲突。
- 2) 两个进程同时执行signal\_sem: 如果signal的信号量相同,信号量的小粒度锁保证了它们不会同时执行;如果signal的信号量不同,由于一个进程不可能同时wait两个信号量,因此这两个进程唤醒的进程必然不同,它们对应的proclist中的元素必然不同,不会发生冲突。
- 3) 两个进程分别执行wait\_sem和signal\_sem:如果操作的信号量相同,信号量的小粒度锁会保证它们不会同时执行;如果操作的信号量不同,由于正在wait的进程不可能正在执行别的代码,因此signal\_sem中唤醒的进程必然不是执行wait\_sem的进程,它们对应的proclist中的元素必然不同,不会发生冲突。

至此,在任何情况下都不会有进程同时读取和修改proclist中的同一个元素,不需要给proclist加锁证毕。

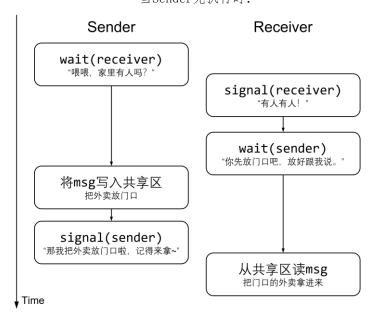
#### 4. 运行结果

## 二、消息传递

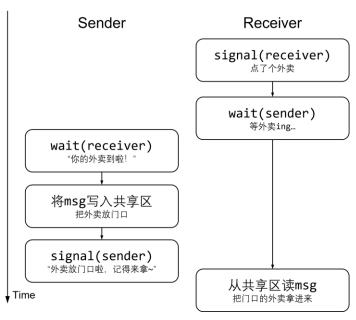
### 1. 思路介绍

本质上,对于每个进程来说,这是多个生产者(消息发送者)和单个消费者(消息接收者),且没有缓冲区(发送者必须等到消息被接受后才能返回)的生产者-消费者问题。因此我参考了多生产者-单消费者的解决方案:使用两个信号量,一个表示receiver是否等待,一个表示正在等待的sender的数量。

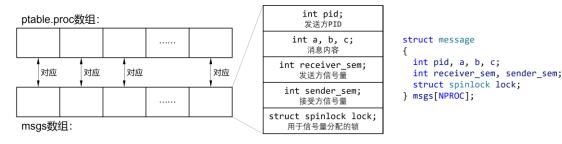
上面这段话可能比较抽象,下面让我用一个"点外卖"的例子来解释: 当Sender先执行时:



当Receiver先执行时:



显然,每个进程都可以作为接收者,因此每个进程都需要两个信号量来控制、和一块 共享内存区用于存放消息内容。和信号量一样,我开辟了一个大小和ptable.proc相同的数 组msgs,其中元素和ptable.proc中的元素按下标一一对应,如下图所示:



事实上,前文所述的send和receive的流程已经可以保证收发双方不会在共享内存区中发生数据竞争,那为什么在message结构体中还需要一个锁呢?这要从信号量的分配说起。

题目要求信号量必须在使用时分配,不能提前分配,因此,我用-1表示信号量没有被分配,在每次wait或signal信号量前检查信号量有没有被分配,如果没有,则立即分配一个。令人遗憾的是,分配信号量的过程可能会发生冲突,即两个sender可能同时判断信号量没有被分配,然后同时分配信号量。所以必须要使用一个锁来保护分配过程。又由于分配信号量的过程很快,使用自旋锁完全可以满足要求。

#### 2. 实现细节

由于很多地方需要检查并分配信号量,我对alloc\_sem做了简单的封装,写了一个异常安全的信号量分配函数:要么分配成功,要么分配失败且没有任何副作用。

另外在进程退出时,即wait函数内,需要释放可能已经分配的信号量。由于dealloc\_sem内部需要获取ptable锁,不能直接在wait函数已经持有ptable锁时调用dealloc\_sem来释放信号量。我的解决方案是: 当在wait函数内调用dealloc\_sem时不让它获取ptable锁,可以通过一个函数参数来控制,具体见代码中的dealloc\_sem\_lk(v, 0)辅助函数。

#### 3. 运行结果

5 msgtest child(ender,child+3): 24(23,77) 25(23,28) 26(23,29) 27(23,28(23,31) 29(23,32) 38(23,33) 31(23,34) 32(23,35) 33(23,36) 34(23,37) 35(23,38) 36(23,39) 38(23,41) 39(23,42) 49(23,43) 41(23,44) 44(23,48) 43(23,46) 37(23,48) 44(23,47) 45(23,48) 46(23,49) 47(23,59) 48(23,51) 49(23,52) 51(23,54) 52(23,55) 53(23,55

### 三、后记

This is not the fastest, most space-conserving, most portable, or most tunable malloc ever written. However it is among the fastest while also being among the most space-conserving, portable and tunable.

Doug Lea

出自https://github.com/lattera/glibc/blob/master/malloc/malloc.c#L46

上文摘自Doug Lea教授为glibc实现的malloc函数中的一段注释,表现了程序设计在时间、空间、可移植性、可微调性等方面的矛盾和取舍。

我在设计信号量时,也经历了这样的取舍。我设计了很多种方案:比如用bitset的思想,当NPROC为默认的64时,可以用一个long long类型记录每个进程是否在等待该信号

量。这样子可以非常节省内存,但是会出现饥饿的情况,可拓展性也有所欠缺。

我最终还是选择了报告中呈现的方案。虽然该方案需要一个看上去莫名其妙的辅助数组,但它可以让alloc、dealloc、wait和signal都在0(1)时间内完成,且不会饥饿。更关键的是,如果进程数量不再受限于NPROC,而是可以动态增长时,这个方案的优点才能真正展现:listNode可以不再与PCB绑定,而是根据实际执行wait sem函数的数量动态分配。