# Операционные Системы Синхронизация потоков

May 5, 2017

## Пример

```
. data
2
3
4
5
    counter:
         .int 0
                                               int counter;
                                          1
2
3
4
                                              int add_one(void)
         . text
6
7
    \mathsf{add}\_\mathsf{one}:
         movq counter, %rax
                                                    return ++counter;
         inc %rax
         movq rax, counter
0
         retq
```

# Вариант 1

movq counter, %rax	1
inc %rax	2
movq rax, counter	3
	4 movq counter, %ra
	5 inc %rax
	6 movg rax, counter

## Вариант 2

```
      1
      movq counter, %rax
      1

      2
      2
      movq counter, %rax

      3
      inc %rax
      3

      4
      movq rax, counter
      4

      5
      inc %rax

      6
      movq rax, counter
```

#### Состояние гонки

- Состояние гонки результат зависит от порядка выполнения инструкций
  - порядок зависит от слишком многих факторов;
  - решения планировщика, влияние других потоков, прерывания...
  - могут быть трудно воспроизводимы не поддаются тестированию.

## Критическая секция

- Критическая секция
  - участок кода, обращающийся к разделяемым несколькими потоками данным;
  - если не более, чем один поток может одновременно находиться в критической секции, то не будет состояния гонки.

## Блокировка

- ▶ Блокировка (lock) некоторый объект и пара методов для работы с ним
  - ▶ lock метод захвата блокировки;
  - unlock метод освобождения блокировки.

- Взаимное исключение (mutual exclusion)
  - потоки всегда вызывают lock и unlock парами (сначала lock, а потом unlock);
  - ▶ не более одного потока может одновременно находиться между lock-ом и unlock-ом.

```
struct lock lock0:
                                struct lock lock1:
int counter0:
                                 int counter1:
int add oneO(void)
                                 int add one1(void)
                                 {
    int res;
                                     int res;
    lock(&lock0);
                                    lock(&lock1);
    res = ++counter0;
                                    res = ++counter1;
    unlock(&lock0);
                             10
                                     unlock(&lock1);
    return res:
                             11
                                     return res:
                             12
                                }
```

#### Свойство живости

#### Свойство живости

- ► Свойство живости (deadlock freedom)
  - если один из потоков вызвал lock, то какой-то из потоков, вызвавших lock, захватит блокировку;
  - поток не ждет в lock, если он единственный пытается захватить блокировку;
  - если поток ждет, значит другому потоку повезло захватить блокировку.

## На что нельзя полагаться?

- Скорость работы потоков:
  - мы не знаем, сколько времени потребуется потоку, чтобы выполнить какой-то код;
  - мы не можем полагать, что какой-то поток быстрее.

## На что можно полагаться?

- Потоки работают корректно:
  - поток не находится между lock и unlock бесконечно;
  - ▶ поток не "падает", находясь между lock и unlock;
  - и так далее...

## Атомарный Read/Write регистр

- Атомарный RW регистр ячейка памяти и пара операций
  - ▶ write "атомарно" записывает значение в регистр;
  - read "атомарно" читает последнее записанное значение;
  - ▶ все операции (read/write) упорядочены.

## Взаимное исключение для 2-х потоков

- Есть всего два потока
  - потоки имеют идентификаторы 0 и 1;
  - внутри потока мы можем узнать его идентификатор (пусть за это отвечает функция threadId).

## Альтернация

```
struct lock {
2
3
4
5
6
7
             atomic int last;
         };
         void lock init(struct lock *lock)
             atomic_store(&lock -> last , 0);
8
9
10
         void lock(struct lock *lock)
11
             while (atomic load(&lock -> last) == threadId());
12
13
14
         void unlock(struct lock *lock)
15
16
17
             atomic_store(&lock->last, threadId());
18
```

- Для приведенного алгоритма взаимное исключение гарантируется
  - ▶ lock может вернуть управление только потоку с идентификатором, не равным lock—>last;
  - ▶ только поток с threadId() != lock->last может изменить значение lock->last.

#### Свойство живости

- ▶ Пусть поток 1 вообще никогда не пытается захватить лок
  - ▶ если поток 0 вызовет lock, то он зависнет навсегда;
  - т. е. свойство живости не выполняется.

## Флаги намерения

```
struct lock {
2
             atomic int flag[2];
3
4
         };
5
         void lock init(struct lock *lock)
6
7
             atomic store(&lock -> flag[0], 0);
8
             atomic store(&lock -> flag[1], 0);
9
         }
10
11
         void lock(struct lock *lock)
12
13
             const int me = threadId();
14
             const int other = 1 - me;
15
             atomic store(&lock -> flag[me], 1);
16
17
             while (atomic load(&lock->flag(other)));
18
19
20
         void unlock(struct lock *lock)
21
22
             const int me = threadId();
23
24
             atomic store(&lock -> flag[me], 0);
25
```

## Корректность

- ▶ Гарантируется ли взаимное исключение?
- ▶ Гарантируется ли живость?

## Алгоритм Петерсона для 2-х потоков

```
struct lock {
             atomic int last;
2
3
4
5
6
7
8
             atomic int flag [2];
         };
         void lock(struct lock *lock)
             const int me = threadId();
9
             const int other = 1 - me;
10
11
             atomic store(&lock -> flag[me], 1);
12
             atomic store(&lock -> last, me);
13
14
             while (atomic load(lock->flag[other])
15
                     && atomic load(&lock -> last) == me);
         }
16
17
18
         void unlock(struct lock *lock)
19
20
             const int me = threadId():
21
22
             atomic store(&lock -> flag[me], 0);
23
```

#### Взаимное исключение

- Доказательство от противного пусть два потока одновременно находятся в критической секции
  - оба потока записывали значение в атомарный регистр last;
  - один из них должен был быть первым, а другой последним;
  - для определенности пусть последним был поток 1.

#### Взаимное исключение

- Итак нам известно следующее:
  - ▶ lock—>last == 1 последним туда записал поток 1;
  - ▶ lock flag[0] = 1 и lock flag[1] == 1.

#### Взаимное исключение

- ▶ Как в таких условиях поток 1 мог пройти мимо цикла в lock и войти в критическую секцию?
  - очевидно, никак.

- Пусть поток 0 пытается войти в критическую секцию, возможны две ситуации:
  - ▶ при проверке условия цикла lock flag[1] == 0;
  - ▶ при проверке условия цикла lock > flag[1] == 1.

- ▶ В первом случае (lock—>flag[1] == 0)
  - поток 1 даже не пытался захватить блокировку;
  - условие цикла, очевидно, ложно, и поток 0 входит в критическую секцию

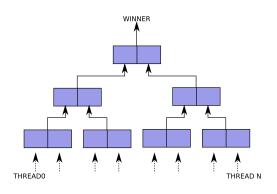
- ▶ Во втором случае (lock -> flag[1] == 1)
  - оба потока изъявили намерение войти в критическую секцию;
  - нужно показать, что хотя бы один из них рано или поздно войдет в критическую секцию (или уже там).

- ▶ Оба потока после записи в lock > flag[x] должны в какой-то момент записать в lock > last
  - ▶ не трудно увидеть, что если lock > flag[0] == 1 и lock > flag[1] == 1,
  - то тот из них, кто сделал это первым, войдет в критическую секцию.

#### N потоков

- Реализовав взаимное исключение для 2-х потоков, мы можем реализовать вазимное исключение для любого числа потоков
  - ▶ организуем турнир для N потоков;
  - потоки конкурируют друг с другом на "выбывание".

### **N** потоков



## Алгоритм Петерсона для N потоков

```
struct lock one {
2
             atomic int last;
             atomic int flag[N];
         };
5
6
         int flags clear(const struct lock one *lock)
7
8
             const int me = threadId();
9
10
             for (int i = 0; i != N; ++i) {
11
                  if (i != me && atomic load(&lock -> flag[i]))
12
                      return 0:
13
14
             return 1;
15
16
17
         void lock one(struct lock one *lock)
18
             const int me = threadId();
19
20
21
             atomic store(&lock -> flag[me], 1);
             atomic store(&lock -> last, me);
22
23
24
             while (!flags clear(lock)
                    && atomic load(&lock->last) == me);
25
26
         }
27
28
         void unlock one(struct lock one *lock)
29
30
             const int me = threadId();
31
32
             atomic store(&lock -> flag[me], 0);
33
         }
```

## Алгоритм Петерсона для N потоков

```
struct lock {
1
2
3
4
5
6
7
8
9
             struct lock one lock [N-1];
         };
         void lock(struct lock *lock)
             for (int i = 0; i != N - 1; ++i)
                  lock one(&lock ->lock[i]);
10
         void unlock(struct lock *lock)
11
12
             for (int i = N - 2; i >= 0; —i)
13
                  unlock one(&lock->lock[i]);
14
15
```

## Алгоритм Петерсона для N потоков

```
struct lock {
2
3
4
5
6
7
8
9
             atomic int level[N];
             atomic int last [N - 1];
         };
         void lock(struct lock *lock)
             const int me = threadId();
10
             for (int i = 0; i != N - 1; ++i) {
                  atomic store(&lock -> level[me], i + 1);
11
12
                  atomic store(&lock -> last[i], me);
13
14
                  while (!flags clear(lock, i)
15
                         && atomic load(&lock -> last[i]) == me);
16
17
18
19
         void unlock(struct lock *lock)
20
21
             const int me = threadId();
22
23
             atomic store(&lock -> level[me], 0);
24
```

#### Честность

- Не хочется, чтобы потоки голодали!
  - если поток захотел захватить блокировку, то когда-нибудь ему это удастся;
  - сравните с живостью среди потоков, пытающихся захватить блокировку, одному это удастся.

## Супер честность

- ▶ k-ограниченное ожидание:
  - после того как поток "изъявил" желание захватить блокировку (встал в очередь), не более k потоков могут пролезть вперед него без очереди.

### Алгоритм Петерсона на примере 3 потоков

Nº	level[0]	level[1]	level[2]	last[0]	last[1]
0	0	0	0	0	0
1	1	0	0	0	0
2	1	1	0	1	0
3	1	1	1	2	0
4	2	1	1	2	0
5	0	1	1	2	0
6	1	1	1	0	0
7	1	1	2	0	2
8	1	1	0	0	2
9	1	1	1	2	2
_10	2	1	1	2	0

# Атомарный Read/Modify/Write регистр

- ▶ Атомарный RMW регистр позволяет за одну операцию
  - прочитать значение в регистре;
  - преобразовать некоторым образом прочитанное значение;
  - записать преобразованное значение назад.

# Атомарный Read/Modify/Write регистр

```
int atomic_rmw(int *reg, int (*f)(int))

const int old = *reg;
const int new = f(old);

*reg = new;
return old;
}
```

# Атомарный Read/Modify/Write регистр

- ▶ atomic\_exchange возвращает старое значение, записывает новое;
- atomic\_fetch\_{add|sub|or|and|xor} выполняет арифметическое дествие над атомарным регистром;
- atomic\_compare\_exchange записывает новое значение, если старое значение равно заданному.

## Реализация RMW регистра

- Архитектура может поддерживать RMW операции (х86 - одна из них)
  - xchg;
  - lock add, lock sub, lock or, lock and, lock xor;
  - lock cmpxchg.

## Реализация RMW регистра

- ▶ Архитектура может поддерживать LL/SC (например, ARM):
  - LL (load-link, load-linked, load-locked) загружает значение из памяти;
  - SC (store-conditional) записывает новое значение в ячейку, но только если после LL эту ячейку никто не трогал;
  - LL/SC идут парами и работают вместе как одна RMW операция.

# Взаимное исключение с использованием RWM регистра

```
#define LOCKED
2
3
4
5
6
7
8
9
        #define UNLOCKED 0
         struct lock {
             atomic int locked;
         };
         void lock(struct lock *lock)
         {
10
             while (atomic exchange(&lock->locked, LOCKED) != UNLOCKED);
11
12
13
         void unlock(struct lock *lock)
14
             atomic store(&lock->locked, UNLOCKED);
15
16
```

### И снова про честность

- ► Что если блокировка находится под нагрузкой (high contention)?
  - т. е. блокировка практически всегда занята;
  - некоторый поток может получать CPU только тогда, когда блокировка занята;
  - ▶ такой поток будет голодать блокировка не честная.

#### Ticket lock

```
struct lock {
2
3
4
5
6
7
8
9
             atomic uint ticket;
             atomic uint next;
         };
         void lock(struct lock *lock)
             const unsigned ticket = atomic_fetch_add(&lock->ticket, 1);
10
             while (atomic load(&lock->next) != ticket);
11
        }
12
13
         void unlock(struct lock *lock)
14
         {
15
             atomic fetch add(&lock->next, 1);
16
```

### И снова о прерываниях

- Пусть у нас есть устройство, которое получает данные из сети
  - устройство сигналит процессору генерирует прерывание;
  - процессор вызывает обработчик прерывания функцию ядра ОС;
  - обработчик прерывания должен забрать данные с устройства и положить их в буфер, из которого какой-то поток сможет их забрать.

#### И снова о прерываниях

- Что если к этому буферу могут обращаться из нескольких потоков?
  - мы должны защитить буфер блокировкой;
  - потоки и обработчики прерываний должны захватывать эту блокировку перед обращением к буферу;
  - что если обработчик прерывания устройства прервал поток, который захватил блокировку?

#### Deadlock

- Прерванный поток и обработчик прерывания ждут друг друга:
  - обработчик прерывания не может захватить блокировку, потому что ее держит прерванный поток;
  - пока обработчик прерывания не завершится, прерванный поток не получит управление и не сможет отпустить блокировку.

### Мораль

- Если блокировка защищает данные, к которым обращается обработчик прерывания, то нужно выключать прерывания
  - если прерывания отключены, то deadlock между обработчиком прерывания и прерванным потоком не может возникнуть.

### Однопроцессорные системы

- Представим систему с всего одним ядром/процессором
  - запретив прерывания и переключение потоков, мы получаем CPU в монопольное пользование;
  - все рассмотренные ранее алгоритмы просто не нужны.

### Разделение на читателей и писателей

- ▶ Не все запросы к разделяемым данным одинаковы
  - есть запросы, которые модифицируют данные;
  - есть запросы, которые только читают данные.

# Разделение на читателей и писателей

```
struct rwlock {
2
3
4
5
6
7
8
             atomic uint ticket;
             atomic uint write;
             atomic uint read;
        };
        void read lock(struct rwlock *lock)
9
             const unsigned ticket = atomic fetch add(&lock->ticket, 1);
10
11
             while (atomic load(&lock->read) != ticket);
12
             atomic store(&lock->read, ticket + 1);
13
14
15
         void read unlock(struct rwlock *lock)
16
17
             atomic_fetch_add(&lock->write, 1);
18
```

# Разделение на читателей и писателей

```
struct rwlock {
    atomic_uint ticket;
    atomic_uint write;
    atomic_uint read;
};

void write_lock(struct rwlock *lock)
{
    const unsigned ticket = atomic_fetch_add(&lock->ticket, 1);
    while (atomic_load(&lock->write) != ticket);
}

void write_unlock(struct rwlock *lock)
{
    atomic_fetch_add(&lock->read, 1);
    atomic_fetch_add(&lock->write, 1);
}
```

## Стратегии ожидания

- ▶ До сих пор функция lock всегда просто ждала в цикле
  - такая стратегия называется активным ожиданием;
  - ▶ блокировки, использующие активное ожидание, часто называются spinlock-ами;
  - они "крутятся" в цикле.

#### Активное ожидание

- Активное ожидание хорошо работает если:
  - потоки не держат блокировку очень долго;
  - блокировка не находится под сильной нагрузкой;
  - т. е. если активное ожидание длится недолго.

#### Альтернативы активному ожиданию

- Как можно ожидать не активно?
  - ▶ можно добровольно отдать CPU (переключиться на другой поток);
  - можно пометить поток как неактивный, чтобы планировщик не давал ему время на CPU, пока блокировка не будет отпущена.

### Задача Producer-а и Consumer-а

- Рассмотрим следующую задачу:
  - Producer поток/потоки, который генерирует данные;
  - Consumer поток/потоки, который потребляет данные;
  - что если Producer и Consumer работают с разной скоростью?

#### Переменная состояния

- ▶ Переменная состояния (condition variable) объект и несколько методов для работы с ним
  - ▶ wait ожидает, пока кто-нибудь не просигналит;
  - notify\_one просигналить одному из ожидающих;
  - notify\_all просигналить всем ожидающим.

#### Переменная состояния

```
struct lock;
void lock(struct lock *lock);

void unlock(struct lock *lock);

struct condition;

void wait(struct condition *cv, struct lock *lock);

void notify_one(struct condition *cv);

void notify_all(struct condition *cv);
```

#### Producer

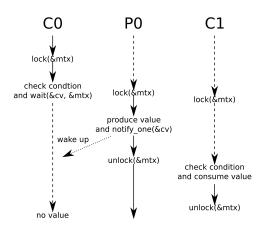
```
struct condition cv;
2
3
4
5
6
7
8
9
         struct lock mtx;
         int value:
         bool valid_value;
         bool done;
         void produce(int x)
              lock(&mtx);
10
              while (valid value)
                  wait(&cv, &mtx);
11
12
             value = x:
13
             valid value = true;
14
              notify one(&cv);
             unlock(&mtx);
15
         }
16
17
18
         void finish (void)
19
20
              lock(&mtx);
21
             done = true;
22
              notify all(&cv);
23
              unlock(&mtx);
24
```

#### Consumer

```
int consume(int *x)
1
2
3
4
5
6
7
8
9
                int ret = 0;
                lock(&mtx);
                while (!valid value && !done)
                      wait(&cv, &mtx);
                if (valid_value) {
    *x = value;
10
                      valid_value = false;
notify one(&cv);
11
12
13
                      ret = 1:
14
15
                unlock(&mtx);
16
                return ret;
17
```

#### Зачем нам lock?

# Зачем нам цикл?



## Зачем нам цикл?

- Spurious wakeups (ложные пробуждения) ситуация, когда wait возвращает управление, даже если никто не сигналил
  - многие реализации переменной состояния подвержены:
    - ► C++;
    - Java;
    - POSIX Threads...

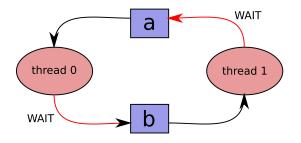
#### Deadlock

- ▶ Deadlock ситуация, при которой потоки не могут работать, потому что ждут друг друга:
  - deadlock потоком исполнения и обработчиком прерывания;
  - поток А ждет, пока поток В что-то сделает (например, отпустит блокировку);
  - а поток В ничего не делает, потому что ждет, пока поток А что-то сделает (например, отпустит блокировку).

```
struct lock b;
   struct lock a;
                                  3
                                     void thread1 (void)
3
   void thread0 (void)
4
                                 5
6
                                          lock(&b);
        lock(&a);
                                          lock(&a);
        lock(&b);
                                          /* do something else
8
        /* do something */
                                                \hookrightarrow */
                                 9
0
        unlock(&b);
                                 10
                                          unlock(&a);
1
        unlock(&b);
                                 11
                                          unlock(&b);
                                 12
                                     }
```

```
1 lock(&a); 1 2 lock(&b); 3 lock(&b); 4 lock(&a);
```

# Wait-for граф



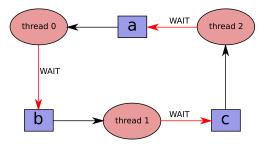
#### Deadlock

- ► Как и с состоянием гонки, deadlock не поддается тестированию
  - появление зависит от многих факторов;
  - входные данные, решения планировщика, прерывания, производительность оборудования ...

## Предотвращение deadlock-ов

- ▶ Мы хотим избежать появления цикла в wait-for графе
  - простой случай все блокировки известны заранее;
  - упорядочим все блокировки (например, по адресу);
  - захватываем блокировки только по порядку.

```
void thread0()
1
2
3
   {
         lock(&a);
        lock(&b);
4
5
6
7
                                   1
2
3
                                       void thread1()
        unlock(&b);
                                       {
        unlock(&a);
                                            lock(&b);
8
                                            lock(&c);
1
   void thread2()
                                            unlock(&c);
2
3
   {
                                            unlock(&b);
         lock(&c);
                                       }
4
5
6
         lock(&a);
        unlock(&a);
7
        unlock(&c);
8
```



- Отсортируем блокировки а, b и с по алфавиту:
  - каждый поток должен захватывать блокировки только согласно порядку;
  - ▶ например, поток 2 хочет захватить блокировки с и а:
    - ▶ так как а в алфавтие раньше с, то сначала хватаем а,
    - потом хватаем с.

```
void thread0()
1
2
3
   {
         lock(&a);
        lock(&b);
4
5
6
7
                                   1
2
3
                                       void thread1()
        unlock(&b);
                                       {
        unlock(&a);
                                            lock(&b);
8
                                            lock(&c);
1
   void thread2()
                                            unlock(&c);
2
3
   {
                                            unlock(&b);
         lock(&a);
                                       }
4
5
6
         lock(&c);
        unlock(&c);
7
        unlock(&a);
8
```

## Предотвращение deadlock-ов

- Сложный случай все блокировки не известны заранее:
  - для этого случая придумано много различных вариантов;
  - ▶ мы рассмотрим подход, который называется Wait-Die.

## Изменим интерфейс

```
struct wdlock_ctx {
    unsigned long long timestamp;
    struct wdlock *next;
};

struct wdlock {
    ...
};

/* Grab unique "timestamp" */
void wdlock_ctx_init(struct wdlock_ctx *ctx);

/* This function may fail */
int wdlock_lock(struct wdlock *lock, struct wdlock_ctx *ctx);

/* Unlocks all of the locks */
void wdlock_unlock(struct wdlock_ctx *ctx);
```

# Как использовать Wait-Die подход?

```
void thread (void)
    struct wdlock ctx ctx;
    wdlock ctx init(&ctx);
    while (1) {
        if (!wdlock lock(&lock1, &ctx)) {
            wdlock unlock(&ctx);
            continue:
        if (!wdlock lock(&lock2, &ctx)) {
            wdlock unlock(&ctx);
            continue;
       Acquired all required locks successfully,
       can do something. */
   wdlock unlock(&ctx);
```

"Контекст"

- ▶ Wait-Die контекст состоит из:
  - списка захваченных блокировок;
  - ▶ уникального "timestamp".

#### "Контекст"

```
struct wdlock ctx {
    unsigned Tong long timestamp;
    struct wdlock *next;
};

void wdlock_ctx_init(struct wdlock_ctx *ctx)

static atomic_ullong timestamp;

ctx_>timestamp = atomic_fetch_add(&timestamp, 1) + 1;
    ctx_>next = NULL;
}
```

### Maruя timestamp

- ▶ timestamp позволяет избегать deadlock-ов
  - храним в каждой блокировке timestamp из wdlock\_ctx, который использовали при захвате блокировки;
  - при попытке захватить блокировку возможно несколько вариантов:
    - если блокировка свободна, то пытаемся ее захватить
       как обычно;
    - если блокировка занята, то нужно сравнить timestamp-ы.

### Maruя timestamp

- ► Если блокировка захвачена, то нужно сравнить наш timestamp с сохраненным в блокировке:
  - если наш timestamp меньше, чем timestamp блокировки, то ждем;
  - в противном случае не ждем, а возвращаем признак неудачи (умираем).

## Корректность

- ► Поток ждет на блокировке, если timestamp блокировки больше, чем timestamp потока
  - deadlock соответствует циклу в Wait-For графе;
  - ▶ при использовании Wait-Die timestamp-ы блокировок на любом пути в графе строго возрастают;
  - следовательно, цикла в Wait-For графе быть не может.

# Wait-Die граф

