Repository Link: <https://github.com/SamanEN/Operating-System-Lab-Projects>

Latest Commit Hash: <hash>

# همگام‌سازی در xv6

## علت غیرفعال شدن وقفه در حین اجرای ناحیه بحرانی، توضیح توابع pushcli و popcli و تفاوت آنها با cli و sti

تابع cli برای غیرفعال کردن وقفه‌ها و تابع sti برای فعال کردن وقفه‌ها استفاده می‌شود. توابع pushcli و popcli به ترتیب، به نوعی یک wrapper برای توابع cli و sti هستند با این تفاوت که می‌توانیم فرض کنیم یک stack مدیریت فعال کردن و یا غیرفعال کردن وقفه‌ها را به عهده می‌گیرد. تا زمانی که این استک خالی است، وقفه‌ها فعال‌اند و به محض اینکه با استفاده از تابع pushcli مقداری در استک push شود، وقفه‌ها غیرفعال می‌شوند. در واقع پیاده‌سازی توابع pushcli و popcli به این شکل است که تابع pushcli به ازای هربار فراخوانی، تابع cli را صدا می‌زند و وقفه‌ها را غیرفعال می‌کند اما تابع popcli تنها زمانی با استفاده از تابع sti وقفه‌ها را فعال می‌کند که استک کاملا خالی باشد. البته لازم به ذکر است که در واقعیت استکی وجود ندارد و فقط تعداد فراخوانی‌های هر یک از توابع در متغیری به نام ncli در هر پردازنده ذخیره می‌شود (به ازای فراخوانی تابع pushcli، مقدار این متغیر یک واحد افزایش پیدا می‌کند و به ازای فراخوانی تابع popcli، مقدار آن یک واحد کاهش می‌یابد) و زمانی که این متغیر برابر با 0 شود، وقفه‌ها فعال می‌شوند و هر موقع مقدار این متغیر بیشتر از 0 شود، وقفه‌ها غیرفعال می‌شوند. پس در واقع 2 بار فراخوانی تابع pushcli، نیازمند 2 بار فراخوانی تابع popcli برای فعال‌سازی مجدد وقفه‌ها است. کاربرد این توابع این است که اگر برای مثال به طور همزمان از دو قفل استفاده می‌کردیم، آزاد کردن یکی از قفل‌ها سبب فعال شدن وقفه‌ها نشود و این مورد فقط زمانی انجام شود که هر دو قفل آزاد شده باشند.

## چرا Spinlock در سیستم‌های تک‌هسته‌ای مناسب نیست؟

در ابتدا نیاز به توضیح مختصری راجع به نحوه عملکرد تابع acquire می‌باشد؛ کد این تابع به شرح زیر است (کامنت‌ها در اینجا حذف شده‌اند):

void  
acquire(struct spinlock \*lk)  
{  
  pushcli();  
  if(holding(lk))  
    panic("acquire");  
  
  while(xchg(&lk->locked, 1) != 0)  
    ;  
  
  \_\_sync\_synchronize();  
  
  lk->cpu = mycpu();  
  getcallerpcs(&lk, lk->pcs);  
}

در این تکه کد، ابتدا تمام وقفه‌ها توسط تابع pushcli غیرفعال می‌شوند. سپس در یک حلقه while دستور اتمی xchg اجرا می‌شود؛ این دستور به صورت اتمی، محتوای یک خانه از حافظه را با یک رجیستر جابجا می‌کند. تابع acquire بصورت مداوم مقدار lk->locked را توسط دستور xchg یک می‌کند؛ در صورتی که قفل از قبل توسط پردازه‌ای دیگر نگه داشته شده باشد، مقدار آن از قبل یک بوده و با اجرای این دستور تفاوتی ایجاد نمی‌شود و تابع xchg مقدار یک را برمی‌گرداند. اما اگر مقدار lk->locked صفر باشد، این دستور مقدارش را یک کرده (که به معنای درگیر بودن قفل است) و تابع xchg مقدار صفر را بازمی‌گرداند؛ در این صورت تابع acquire از حلقه خارج شده و پردازه ادامه روند اجرایش را از سر می‌گیرد.

واضح است که این روش باعث انتظار مشغول[[1]](#footnote-1) می‌شود. انتظار مشغول در سیستم‌های چند پردازنده‌ای باعث هدر رفتن زمان پردازنده و در نتیجه افت بهینگی سیستم شود؛ اما در سیستم‌های تک پردازنده‌ای این موضوع در بدترین حالت می‌تواند منجر به deadlock شود. حالتی را فرض کنید که در آن یک پردازه قفلی را در اختیار می‌گیرد؛ سپس پردازه‌ای دیگر سعی می‌کند قفل را به روش مذکور بدست بیاورد؛ در این صورت پردازه دوم هیچگاه از حلقه خارج نشده و پردازه‌‌های دیگر زمان‌بندی نمی‌شوند.

## مختصری راجع به تعامل میان پردازه‌ها توسط توابع Sleeplock و چرا استفاده از Spinlock در مثال producer-consumer ممکن نیست؟

یک sleeplock در xv6 از اجزای زیر تشکیل شده است:

struct sleeplock {  
 uint locked;  
 struct spinlock lk;  
 char \*name;  
 int pid;  
};

متغیر locked وضعیت قفل بودن را نشان داده و از یک spinlock برای حفاظت از کل اجزای استراکت استفاده می‌شود.

وقتی یک پردازه acquiresleep می‌کند، طبق محتوای تابع:

void  
acquiresleep(struct sleeplock \*lk)  
{  
 acquire(&lk->lk);  
 while (lk->locked) {  
 sleep(lk, &lk->lk);  
 }  
 lk->locked = 1;  
 lk->pid = myproc()->pid;  
 release(&lk->lk);  
}

ابتدا spinlock گرفته شده و سپس بررسی می‌کند که آیا sleeplock در دست پردازه دیگریست یا خیر.

در صورت آزاد بودن لاک، خود پردازه آن را گرفته و در غیر این صورت، پردازه sleep می‌شود.

نسخه ساده شده تابع sleep:

void  
sleep(void \*chan, struct spinlock \*lk)  
{  
 struct proc \*p = myproc();  
 if(p == 0 || lk ==0)  
 panic("...");  
  
 release(lk);  
 p->chan = chan;  
 p->state = SLEEPING;  
  
 sched();  
  
 p->chan = 0;  
 acquire(lk);  
}

رفتار این تابع به این شکل است که از یک spinlock (که اینجا همان spinlock داخل sleeplock است) که از قبل گرفته شده را دریافت کرده و قبل از تغییر وضعیت پردازه به SLEEPING، آن را release می‌کند. بعداً پس از wakeup شدن و بازگشت به این بخش از طرف scheduler، لاک دوباره acquire می‌شود.

این تابع یک آرگومان void\* chan نیز دریافت می‌کند. این متغیر در struct proc هم وجود داشته و در sleep آن را به ورودی تغییر می‌دهد. اینجا خود sleeplock به عنوان chan پاس داده شده است.

حال که پردازه در حالت اجرا شونده نیست، پردازه‌ای که قفل را داشت releasesleep می‌کند:

void  
releasesleep(struct sleeplock \*lk)  
{  
 acquire(&lk->lk);  
 lk->locked = 0;  
 lk->pid = 0;  
 wakeup(lk);  
 release(&lk->lk);  
}

وقتی که پردازه قبلی sleep کرد، لاک release شد پس اینجا می‌تواند آن را acquire کند. پردازه قفل را رها کرده و wakeup را روی خود متغیر sleeplock صدا می‌زند.

تابع wakeup یک void\* chan را دریافت کرده و با پیمایش روی همه پردازه‌ها، وضعیت آنهایی که روی chan خوابیده اند را به RUNNABLE تغییر می‌دهد تا بعدا توسط scheduler قابلیت انتخاب شدن برای اجرا داشته باشند.

از آنجا که پردازه‌ای که برای انتظار دریافت sleeplock روی خود لاک sleep کرده است، اینجا هم با صدا زدن wakeup روی خود لاک بیدار می‌شود.

دلیل اینکه در acquiresleep خود sleep کردن در یک چرخه است، برای این است که اگر به دلیلی wakeup شد ولی هنوز لاک قفل بود، به sleep بازگردد تا releasesleep صدا شود.

## توضیح حالات مختلف پردازه‌ها در xv6 و وظیفه تابع sched()

حالت یک پردازه در متغیر state آن که از جنس enum procstate است نگه داشته می‌شود:

enum procstate { UNUSED, EMBRYO, SLEEPING, RUNNABLE, RUNNING, ZOMBIE };

* UNUSED: از آنجا که پردازه‌ها در یک لیست 64 تایی نگه داشته می‌شوند، خانه‌هایی که درشان پردازه حقیقی ای نیست با این state مشخص شده‌اند و به معنی نبود پردازه است.
* EMBRYO: وقتی که پردازه جدید ساخته می‌شود (مثلا با fork) در ابتدا حالت پردازه این است. یعنی تابع allocproc از بین پردازه‌های UNUSED یکی را انتخاب و آن را EMBRYO می‌کند.
* SLEEPING: در این وضعیت، پردازه در بین انتخاب‌های scheduler برای تخصیص پردازنده به آن قرار نمی‌گیرد و بدون هیچ فعالیتی می‌ماند. پردازه می‌تواند به صورت داوطلبانه یا توسط کرنل به این حالت برود و در انتظار دسترسی به یک منبع بماند.
* RUNNABLE: وقتی پردازه در این حالت است، یعنی در صف اجرای scheduler قرار دارد و در یکی از راندهای زمان‌بندی بعدی CPU به آن می‌رسد و RUNNING می‌شود. چند حالت که منجر به ورود به این وضعیت می‌شوند:
  + پردازه تازه تشکیل شده و از EMBRYO به RUNNABLE میاید.
  + پردازه در حال اجرا و RUNNING بوده و با اتمام time slice یا yield توسط کرنل پردازنده از آن گرفته می‌شود.
  + پردازه در SLEEPING بوده و با wakeup قابل اجرا می‌شود.
  + پردازه‌ای که SLEEPING بوده kill شده و پس از 1 کردن فیلد killed آن پردازه، به حالت RUNNABLE میاید تا وقتی دوباره اجرا می‌شود، همان اول به کشته شدن پی برده شده و exit شود.
* RUNNING: این حالت یعنی پردازه درحال اجرا توسط پردازنده است. تعداد پردازه‌های RUNNING در یک زمان حداکثر معادل تعداد پردازنده‌ها است.
* ZOMBIE: وقتی پردازه کارش تمام می‌شود و می‌خواهد exit بکند، ابتدا ZOMBIE می‌شود. یعنی مستقیم به UNUSED نمی‌رود و در حالتی می‌ماند تا پدرش بتواند خبر بیابد (با wait کردن) که کار فرزندش تمام شده.

تابع sched() برای context switch کردن به context زمان‌بند است. پردازه برای رها کردن CPU به این تابع میاید (که از قبل باید state اش از RUNNING عوض شده باشد و قفل ptable را داشته باشد). در تابع فلگ interrupt enable ذخیره شده و پس از بازگشت برگردانده می‌شود.

این تابع با استفاده از swtch که در آزمایش سوم کامل توضیح داده شده و در اسمبلی نوشته شده، context را تغییر می‌دهد و ادامه تابع scheduler اجرا می‌شود که به context پردازه RUNNABLE دیگری تعویض می‌کند.

## تغییری در توابع Sleeplock بدهید تا تنها پردازه صاحب قفل، قادر به آزادسازی آن باشد و قفل معادل در هسته لینوکس را به طور مختصر معرفی کنید

در استراکت sleeplock متغیر pid قرار داده شده که منظور اولیه از ایجاد آن برای دیباگ کردن بوده. از همین متغیر برای چک کردن صاحب قفل در releasesleep استفاده می‌کنیم. تکه کد زیر، تابع جدید است که قابلیت خواسته شده در آن ایجاد شده:

void  
releasesleep(struct sleeplock \*lk)  
{  
  acquire(&lk->lk);  
  
  if (lk->pid != myproc()->pid) {  
    release(&lk->lk);  
    return;  
  }  
  
  lk->locked = 0;  
  lk->pid = 0;  
  wakeup(lk);  
  release(&lk->lk);  
}

در هسته لینوکس از قفل mutex استفاده می‌شود. تعریف این قفل بصورت زیر می‌باشد (در فایل [mutex.h](https://github.com/torvalds/linux/blob/master/include/linux/mutex.h) موجود است):

    struct mutex {  
        atomic\_long\_t owner;  
        raw\_spinlock\_t wait\_lock;  
#ifdef CONFIG\_MUTEX\_SPIN\_ON\_OWNER  
        struct optimistic\_spin\_queue osq; */\* Spinner MCS lock \*/*  
#endif  
        struct list\_head wait\_list;  
#ifdef CONFIG\_DEBUG\_MUTEXES  
        void\* magic;  
#endif  
#ifdef CONFIG\_DEBUG\_LOCK\_ALLOC  
        struct lockdep\_map dep\_map;  
#endif  
    };

همانطور که قابل مشاهده است، یک فیلد به نام owner برای این استراکت در نظر گرفته شده است؛ این فیلد در حین رها کردن قفل چک می‌شود تا تنها صاحب قفل مجاز به این کار باشد.

## افزونه‌های همگام‌سازی تراکنشی (TSX) را شرح دهید و نقش حذف قفل را در آن بیان کنید

ایده transactional memory به این صورت است که تراکنش‌های حافظه داریم. هر تراکنش یک دنباله از خواندن و نوشتن‌ها در حافظه است که باید به صورت اتمی و پشت سر هم انجام بپذیرد.

در صورتی که تراکنش به صورت کامل انجام شود، تغییرات اعمال شده و در غیر این صورت، عملیات متوقف شده و اعمال نمی‌شوند.

روش‌های لاک عادی وقتی contention زیاد می‌شود (بیشتر شدن پردازنده‌ها هم به زیاد شدن contention می‌تواند منجر شود) مشکلاتی از جمله ددلاک و کند شدن را به وجود میاورد. برای همین به دنبال روش‌های دیگری برای هندل کردن critical section هستیم.

اگر در زبان برنامه‌نویسی بلاک کدی وجود داشته باشد که به کامپایلر بگوید که این بخش باید به صورت تراکنش انجام بپذیرد، از مزایای این روش می‌توان بهره‌مند شد؛ چون در این صورت، برنامه نویس دیگر درگیر لاک‌ها نمی‌شود و اتمی بودن عملیات در پشت صحنه هندل می‌شود.

حافظه تراکنشی در دو سطح نرم‌افزار و سخت‌افزار قابل پیاده شدن است. STM (Software TM) کاملا در نرم‌افزار بوده و کامپایلر کد را بررسی کرده و کدهای مورد نیاز را وارد می‌کند. HTM (Hardware TM) تدخلات داده‌های اشتراک گذاشته شده را در سمت سخت افزار حل می‌کند و مشکل کند بودن STM را ندارد.

TSX (Transactional Synchronization Extentions) یک افزونه برای ISA اینتل x86 است که ساپورت HTM را اضافه می‌کند. این HTM سرعت اجرای برنامه‌های مولتی ترد را با استفاده از lock elision بهبود می‌دهد. این افزونه باگ‌هایی داشته و در خیلی از پردازنده‌های اینتل که قابلیت اش را داشتند غیر فعال شده است. برای همین برنامه‌ها هنوز خیلی از TSX استفاده نمی‌کنند.

در این طراحی، با استفاده از اینستراکشن‌های جدید، سعی بر تراکنش می‌شود و اگر در زمان انجام نشد و abort شد، به لاک معمولی بر می‌گردد که به این کار lock elision می‌گویند.

یعنی با استفاده از این روش، ظاهر کد یکسان می‌ماند و لاک‌ها مثل یک wrapper اند که ابتدا سعی بر تراکنش می‌کنند و در صورت جواب ندادن به لاک عادی بر می‌گردند. و این یعنی لایبرری ارائه‌دهنده لاک‌ها باید تغییرات استفاده از TSX را اعمال کند.

# مانع

## پیاده‌سازی ماکروی barrier() در لینوکس برای معماری x86

#define barrier() \_\_asm\_\_ \_\_volatile\_\_("": : :"memory")

## آیا یک دستور مانع حافظه باید مانع بهینه‌سازی هم باشد؟ نام ماکرو پیاده‌سازی سه نوع مانع حافظه در لینوکس برای معماری x86 را به همراه دستورهای ماشین پیاده‌سازی آنها ذکر کنید

بله یک مانع حافظه باید یک مانع بهینه‌سازی نیز باشد زیرا پردازنده فقط برنامه کامپایل شده را می‌بیند و ترتیت خطوط کدها پیش از بهینه‌سازی را نمی‌داند. برای مثال شبه کد زیر را در نظر بگیرید:

Statement1;

Statement2;

MemBarrier;

Statement3;

اگر فرض کنیم Statement2 و Statement3 به همدیگر وابسته نباشند، کامپایلر در هنگام بهینه‌سازی ممکن است ترتیب این 2 دستور را عوض کند. اما پردازنده به دلیل وجود مانع، تغییری در ترتیب دستورات نمی‌دهد و دستورات به ترتیب زیر اجرا می‌شوند:

Statement1;

Statement3;

Statement2;

در نتیجه ترتیب دستورات برخلاف خواسته توسعه‌دهنده، تغییر کرده است.

در لینوکس به طور کلی 3 دستور برای ایجاد مانع حافظه وجود دارد که در بخش زیر توضیح داده شده‌اند:

1. دستور rmb: این دستور از انتقال هرگونه دستور read access به سمت دیگر مانع، جلوگیری می‌کند. دستورالعمل ماشین آن در x86، lfence است که حرف l در آن، نشان‌دهنده کلمه load است.
2. دستور wmb: این دستور از انتقال هرگونه دستور write access به سمت دیگر مانع جلوگیری می‌کند. دستورالعمل ماشین آن در x86، sfence است که حرف s در آن، نشان‌دهنده کلمه store است.
3. دستور mb: این دستور از انتقال هرگونه دستور memory access به سمت دیگر مانع جلوگیری می‌کند (در واقع ترکیبی از 2 دستور قبلی است). دستورالعمل ماشین آن در x86، mfence است که حرف m در آن، نشان‌دهنده کلمه memory است.

برای یافتن دستورات مانع حافظه در لینوکس و دستورالعمل ماشین معادل آن‌ها در معماری x86، از فایل [linux/arch/x86/include/asm/barrier.h](https://github.com/torvalds/linux/blob/master/arch/x86/include/asm/barrier.h) استفاده شده است که کد آن در بخش زیر قابل مشاهده است:

#ifdef CONFIG\_X86\_32  
#define mb() asm volatile(ALTERNATIVE("lock; addl $0,-4(%%esp)", "mfence", \  
 X86\_FEATURE\_XMM2) ::: "memory", "cc")  
#define rmb() asm volatile(ALTERNATIVE("lock; addl $0,-4(%%esp)", "lfence", \  
 X86\_FEATURE\_XMM2) ::: "memory", "cc")  
#define wmb() asm volatile(ALTERNATIVE("lock; addl $0,-4(%%esp)", "sfence", \  
 X86\_FEATURE\_XMM2) ::: "memory", "cc")  
#else  
#define \_\_mb() asm volatile("mfence":::"memory")  
#define \_\_rmb() asm volatile("lfence":::"memory")  
#define \_\_wmb() asm volatile("sfence" ::: "memory")  
#endif

در لینوکس توابع دیگری نیز برای ایجاد مانع وجود دارند. برای مثال ماکرو barrier یک مانع نرم‌افزاری (بهینه‌سازی) به وجود می‌آورد. ماکروهای smp\_rmb، smp\_wmb و smp\_mb نیز در سیستم‌های چندهسته‌ای به ترتیب معادل دستورات 1 تا 3 هست و در سیستم‌های تک هسته‌ای معادل دستور barrier هستند.

## کابرد مانع در پردازش موازی

# پیاده‌سازی Semaphore

# شبیه‌سازی مسئله فلاسفه خورنده

1. Busy waiting [↑](#footnote-ref-1)