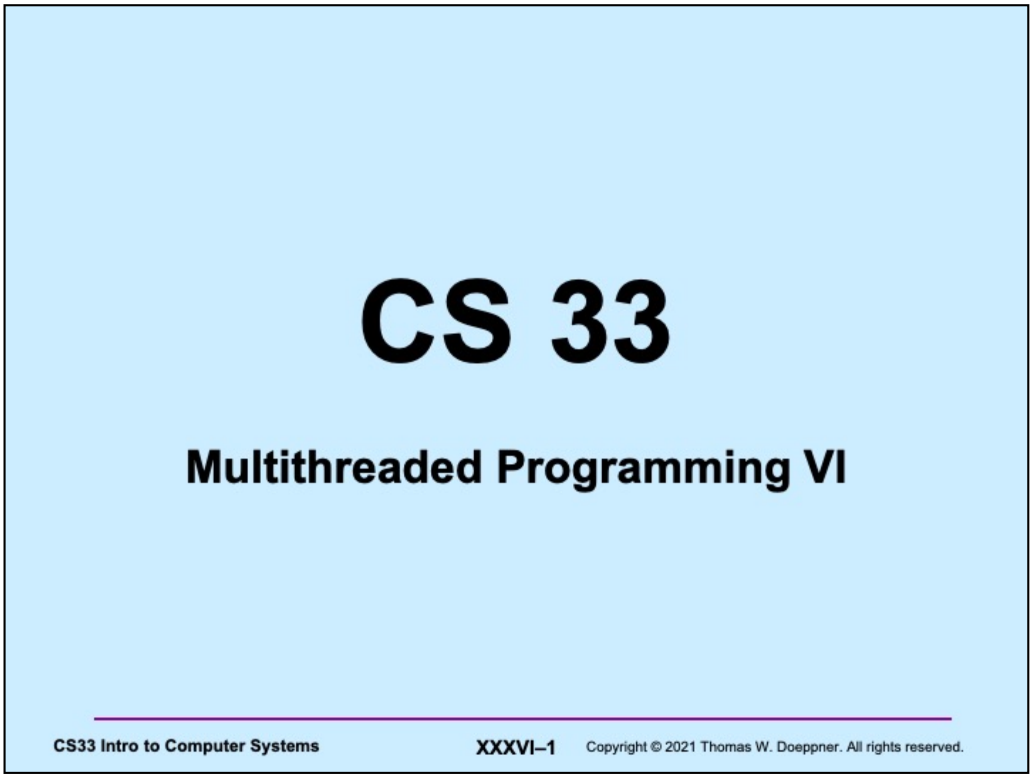
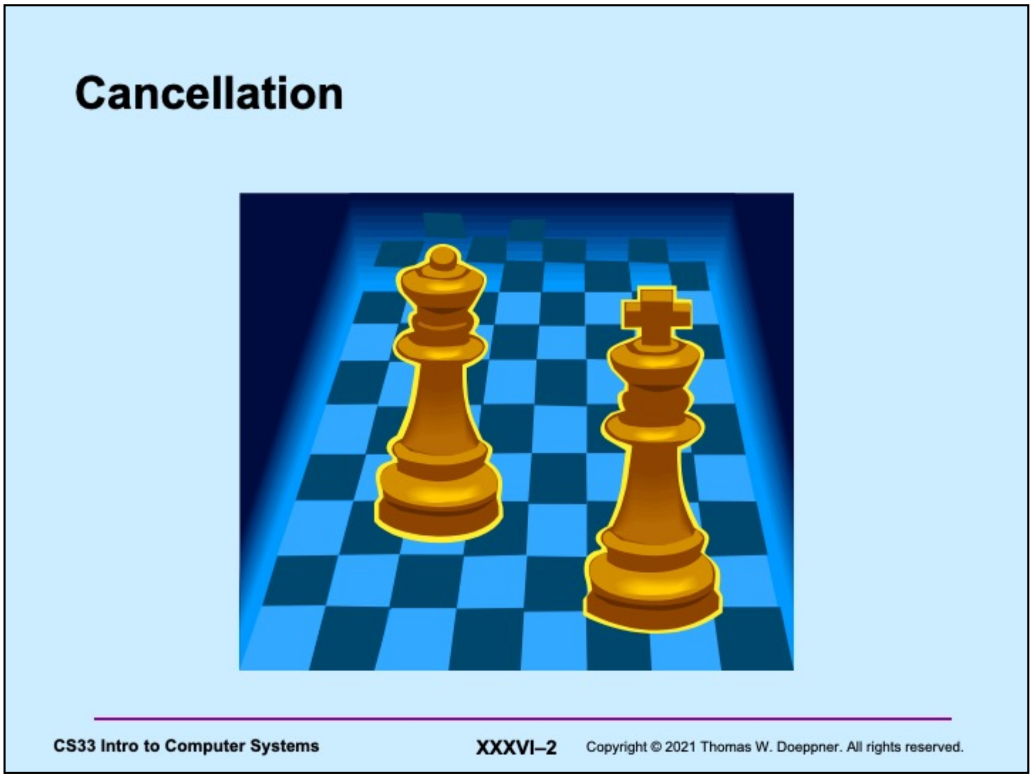


CS 33

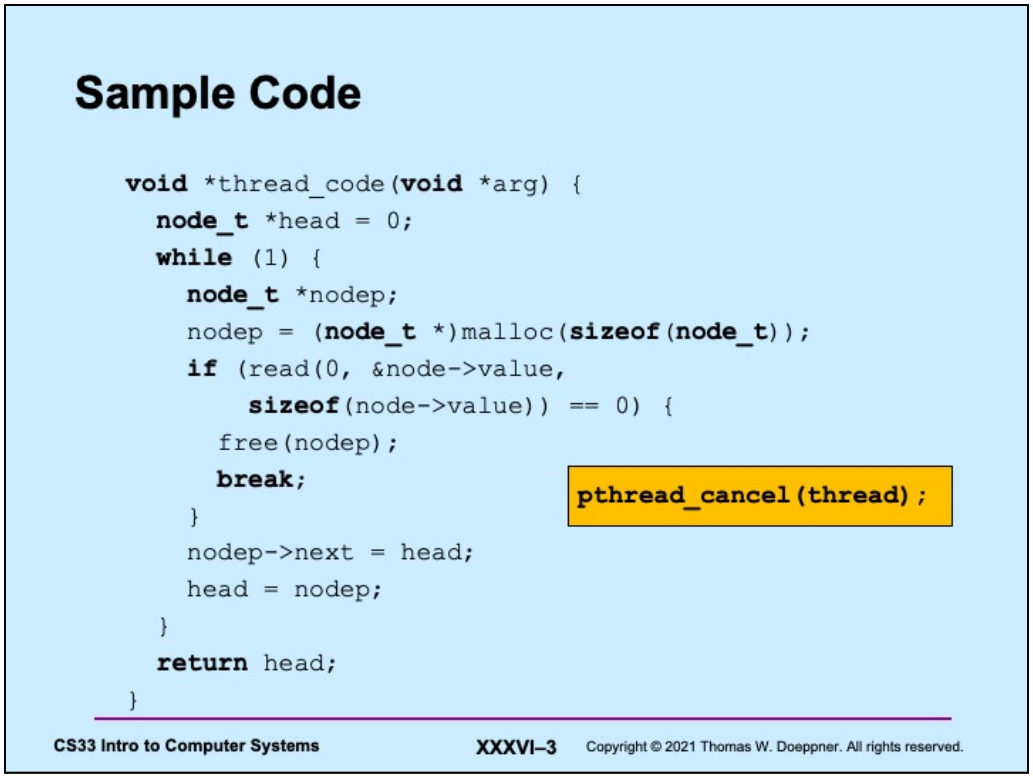
Multithreaded Programming VI



Cancellation

In a number of situations one thread must tell another to cease whatever it is doing. For example, suppose we've implemented a chess-playing program by having multiple threads search the solution space for the next move. If one thread has discovered a quick way of achieving a checkmate, it would want to notify the others that they should stop what they're doing, the game has been won.

One might think that this is an ideal use for per-thread signals, but there's a cleaner mechanism for doing this sort of thing in POSIX threads, called **cancellation**.



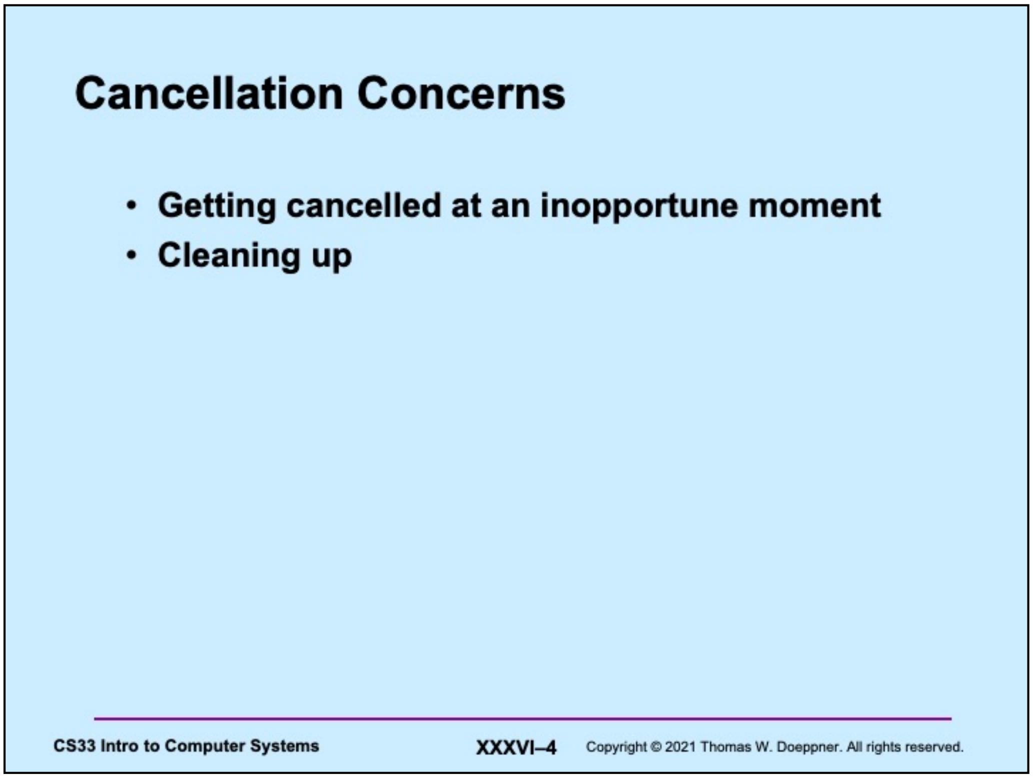
Sample Code

```
void *thread_code(void *arg) {
  node_t *head = 0;
  while (1) {
    node_t *nodep;
    nodep = (node_t *)malloc(sizeof(node_t));
    if (read(0, &node->value,
        sizeof(node->value)) == 0) {
      free(nodep);
      break;
    }
    nodep->next = head;
    head = nodep;
  }
  return head;
}
```

```
pthread_cancel(thread);
```

This code is invoked by a thread (as its first function). The thread reads values from stdin, which it then puts into a singly linked list that it allocates on the fly, and returns a pointer to the list.

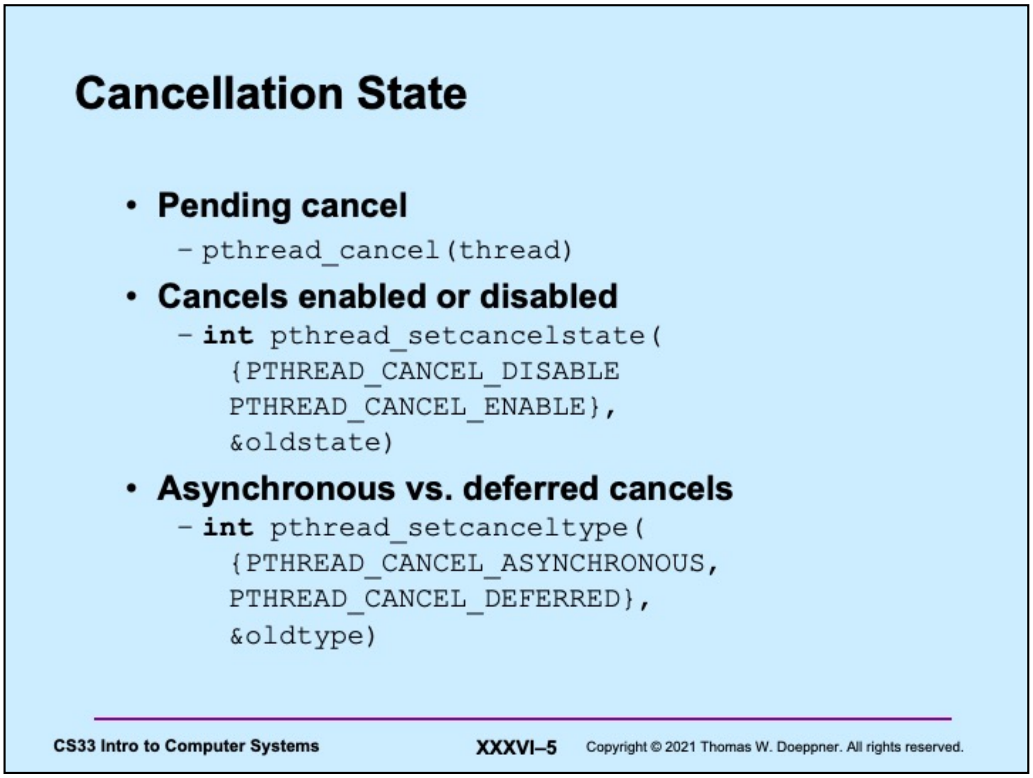
Suppose our thread is forced to terminate in the midst of its execution (some other thread invokes the operation **pthread_cancel** on it). What sort of problems might ensue?



Cancellation Concerns

- Getting cancelled at an inopportune moment
- Cleaning up

We have two concerns about the forced termination of threads resulting from cancellation: a thread might be in the middle of doing something important that it must complete before self-destructing; and a canceled thread must be given the opportunity to clean up.



Cancellation State

- **Pending cancel**
  - `pthread_cancel(thread)`
- **Cancels enabled or disabled**

```
– int pthread_setcancelstate(
    {PTHREAD_CANCEL_DISABLE
    PTHREAD_CANCEL_ENABLE},
    &oldstate)
```

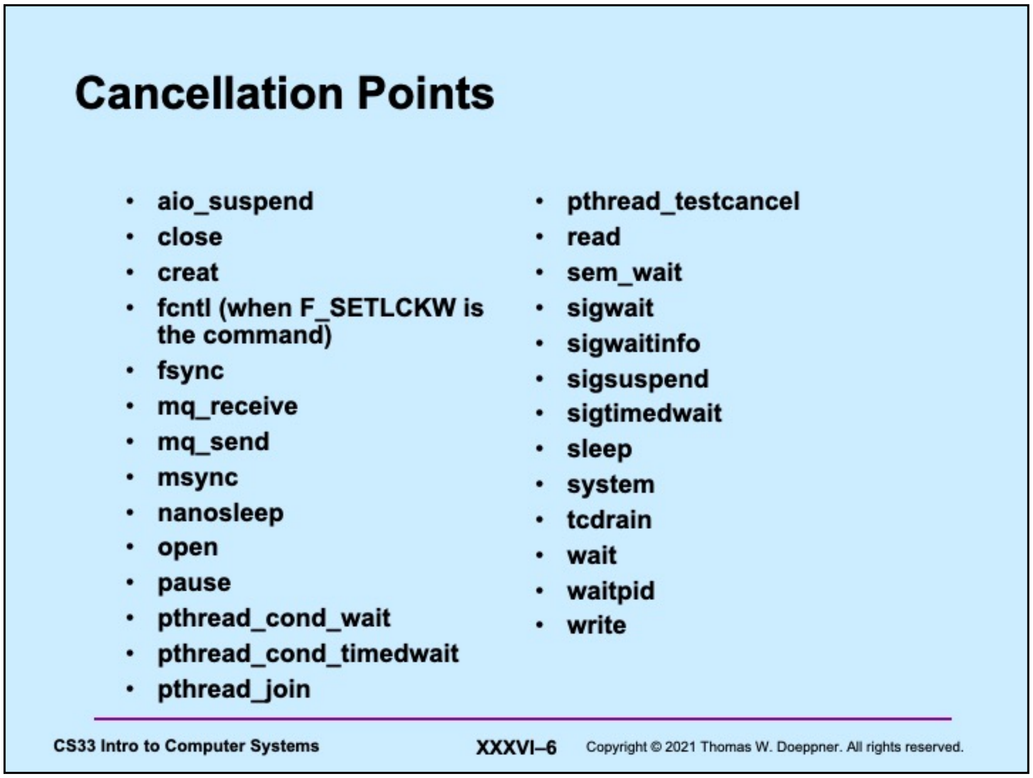
- **Asynchronous vs. deferred cancels**

```
– int pthread_setcanceltype(
    {PTHREAD_CANCEL_ASYNCHRONOUS,
    PTHREAD_CANCEL_DEFERRED},
    &oldtype)
```

A thread issues a cancel request by calling **pthread_cancel**, supplying the ID of the target thread as the argument. Associated with each thread is some state information known as its **cancellation state** and its **cancellation type**. When a thread receives a cancel request, it is marked indicating that it has a pending cancel. The next issue is when the thread should notice and act upon the cancel. This is governed by the cancellation state: whether cancels are **enabled** or **disabled** and by the cancellation type: whether the response to cancels is **asynchronous** or **deferred**. If cancels are **disabled**, then the cancel remains pending but is otherwise ignored until cancels are enabled. If cancels are **enabled**, they are acted on as soon as they are noticed if the cancellation type is **asynchronous**. Otherwise, i.e., if the cancellation type is *deferred*, the cancel is acted upon only when the thread reaches a **cancellation point**.

Cancellation points are intended to be well defined points in a thread's execution at which it is prepared to be canceled. They include pretty much all system and library calls in which the thread can block, with the exception of **pthread_mutex_lock**. In addition, a thread may call **pthread_testcancel**, which has no function other than being a cancellation point.

The default is that cancels are enabled and deferred. One can change the cancellation state of a thread by using the routines shown in the slide. Calls to **pthread_setcancelstate** and **pthread_setcanceltype** return the previous value of the affected portion of the cancellability state.

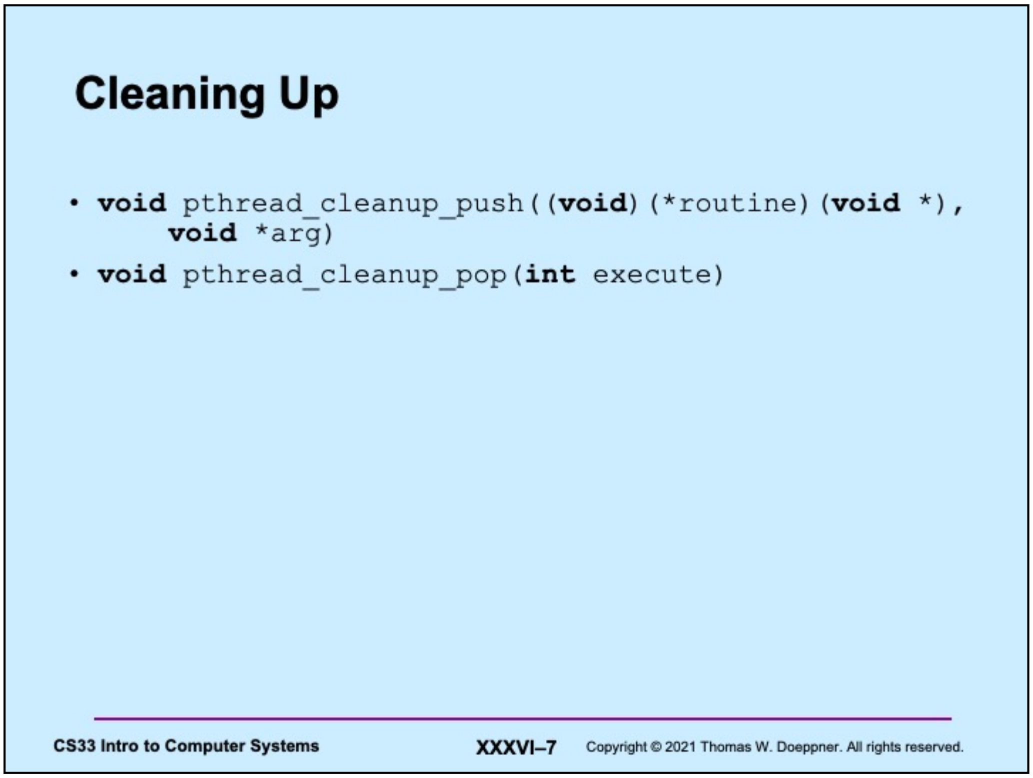


Cancellation Points

- aio_suspend
- close
- creat
- fcntl (when F_SETLCKW is the command)
- fsync
- mq_receive
- mq_send
- msync
- nanosleep
- open
- pause
- pthread_cond_wait
- pthread_cond_timedwait
- pthread_join
- pthread_testcancel
- read
- sem_wait
- sigwait
- sigwaitinfo
- sigsuspend
- sigtimedwait
- sleep
- system
- tcdrain
- wait
- waitpid
- write

The slide lists all of the required cancellation points in POSIX.

The function **pthread_testcancel** is strictly a cancellation point — it has no other function. If there are no pending cancels when it is called, it does nothing and simply returns.

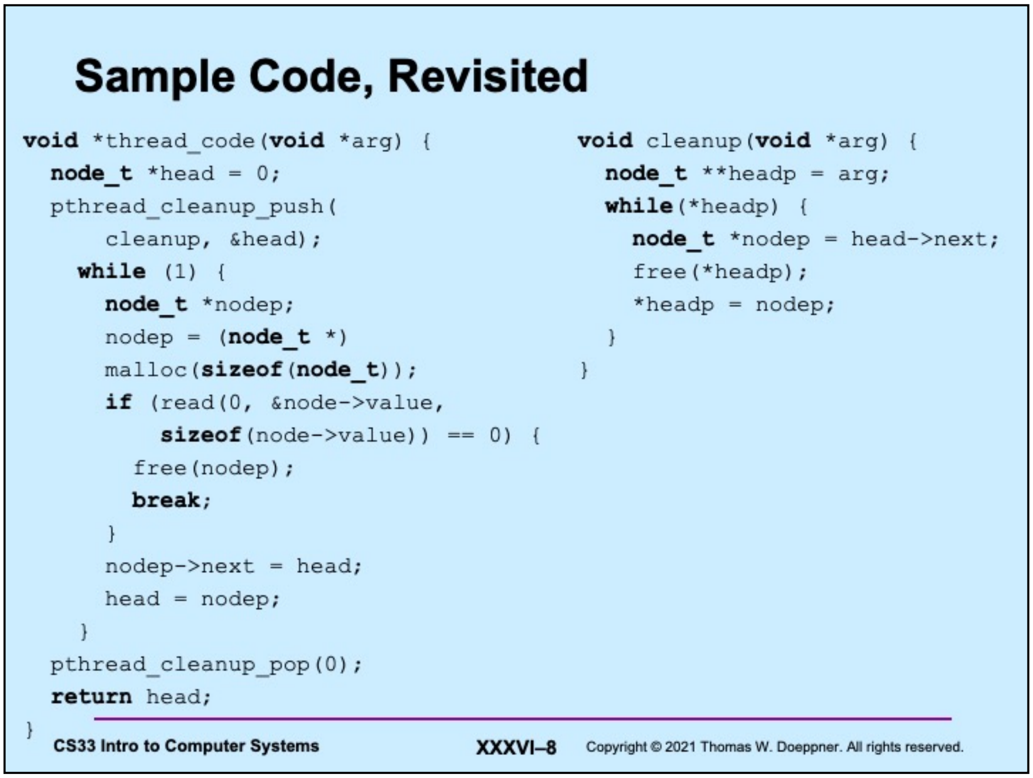


Cleaning Up

- `void pthread_cleanup_push((void)(*routine)(void *),`
  `void *arg)`
- `void pthread_cleanup_pop(int execute)`

When a thread acts upon a cancel, its ultimate fate has been established, but it first gets a chance to clean up. Associated with each thread may be a stack of *cleanup handlers*. Such handlers are pushed onto the stack via calls to **pthread_cleanup_push** and popped off the stack via calls to **pthread_cleanup_pop**. Thus, when a thread acts on a cancel or when it calls **pthread_exit**, it calls each of the cleanup handlers in turn, giving the argument that was supplied as the second parameter of **pthread_cleanup_push**. Once all the cleanup handlers have been called, the thread terminates.

The two functions **pthread_cleanup_push** and **pthread_cleanup_pop** are intended to act as left and right parentheses, and thus should always be paired (in fact, they may actually be implemented as macros: the former contains an unmatched “{“, the latter an unmatched “}”). The argument to the latter function indicates whether or not the cleanup function should be called as a side effect of calling **pthread_cleanup_pop**.

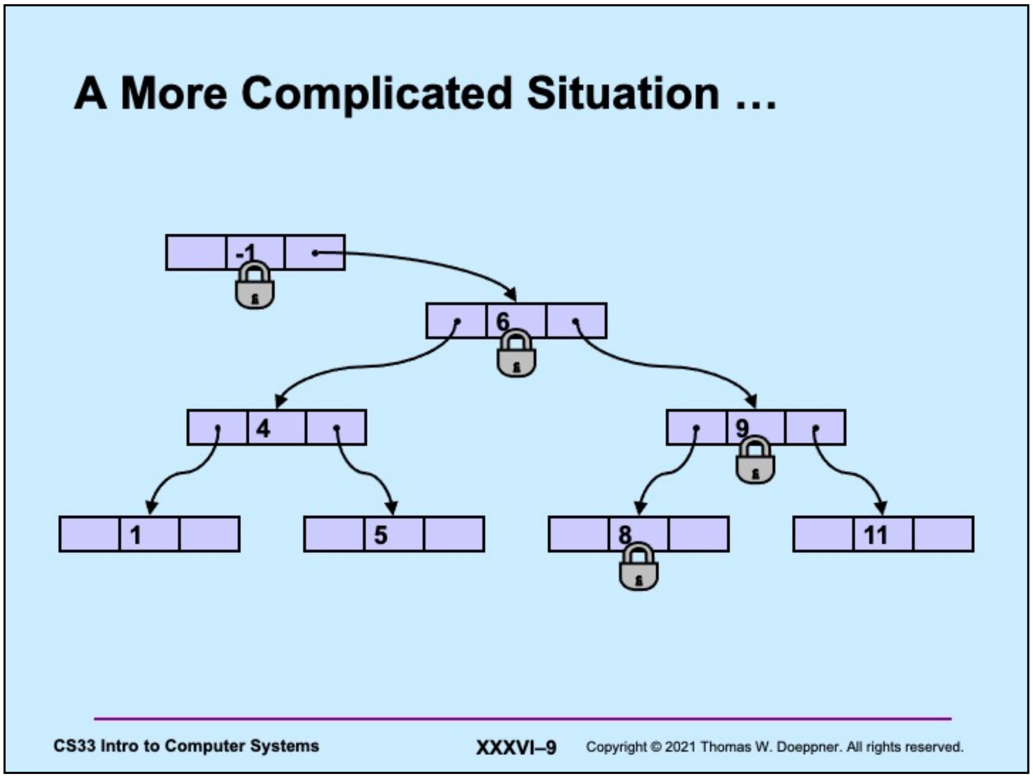


Sample Code, Revisited

```
void *thread_code(void *arg) {
  node_t *head = 0;
  pthread_cleanup_push(
      cleanup, &head);
    while (1) {
      node_t *nodep;
      nodep = (node_t *)
      malloc(sizeof(node_t));
      if (read(0, &node->value,
          sizeof(node->value)) == 0) {
        free(nodep);
        break;
      }
      nodep->next = head;
      head = nodep;
    }
  pthread_cleanup_pop(0);
  return head;
}
```

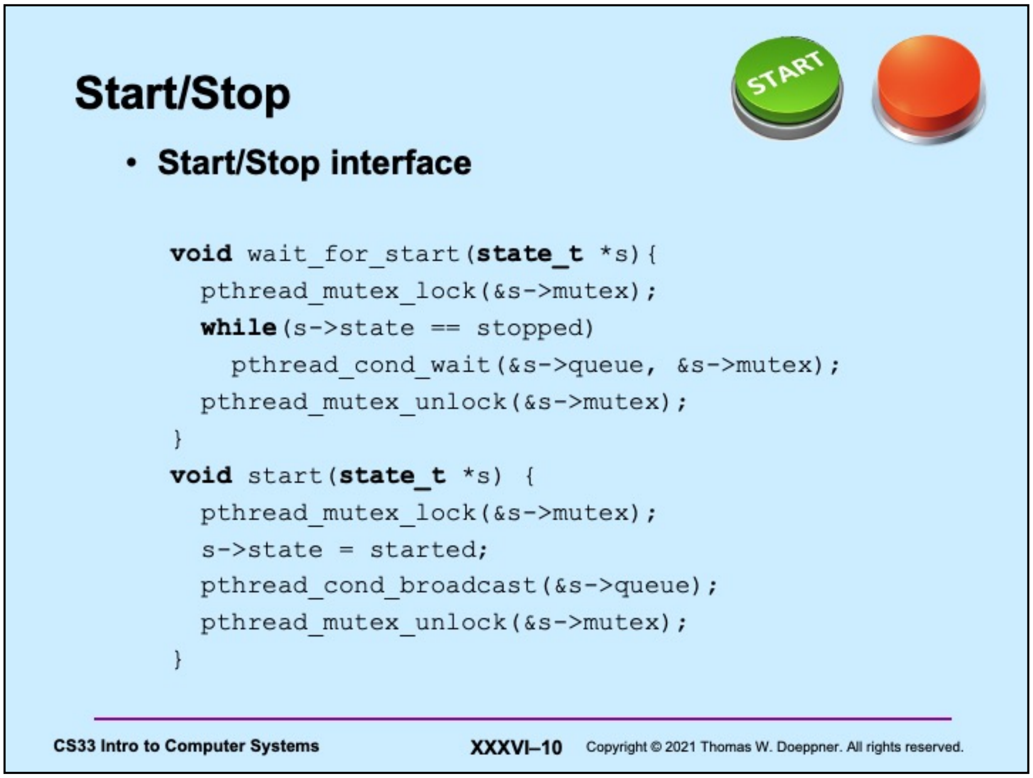
```
void cleanup(void *arg) {
  node_t **headp = arg;
  while(*headp) {
    node_t *nodep = head->next;
    free(*headp);
    *headp = nodep;
  }
}
```

Here we've added a cleanup handler to our sample code. Note that our example has just one cancellation point: **read**. The cleanup handler iterates through the list, deleting each element.



A More Complicated Situation …

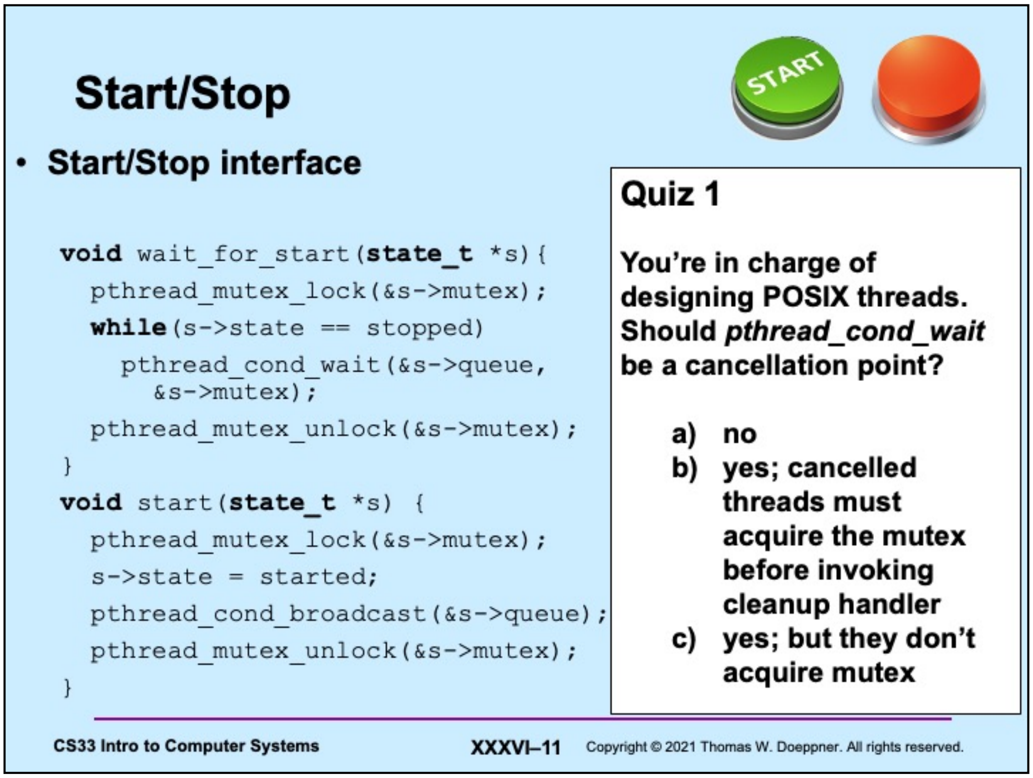
Whether threads are using mutexes or readers/writers locks when manipulating a search tree, if we have to deal with cancellation points in the middle of such operations, things can get pretty complicated and error-prone. Thus, the operations to lock mutexes and readers/writers locks are not cancellation points. (Note, however, that for the case of readers/writers locks, POSIX permits waiting for readers/writers locks to be cancellation points, for the sake of vendors who have poor implementations of them. Neither Linux nor OSX implements such waiting as cancellation points.)



Start/Stop

- **Start/Stop interface**

```
void wait_for_start(state_t *s){
  pthread_mutex_lock(&s->mutex);
  while(s->state == stopped)
    pthread_cond_wait(&s->queue, &s->mutex);
  pthread_mutex_unlock(&s->mutex);
}
void start(state_t *s) {
  pthread_mutex_lock(&s->mutex);
  s->state = started;
  pthread_cond_broadcast(&s->queue);
  pthread_mutex_unlock(&s->mutex);
}
```



Start/Stop

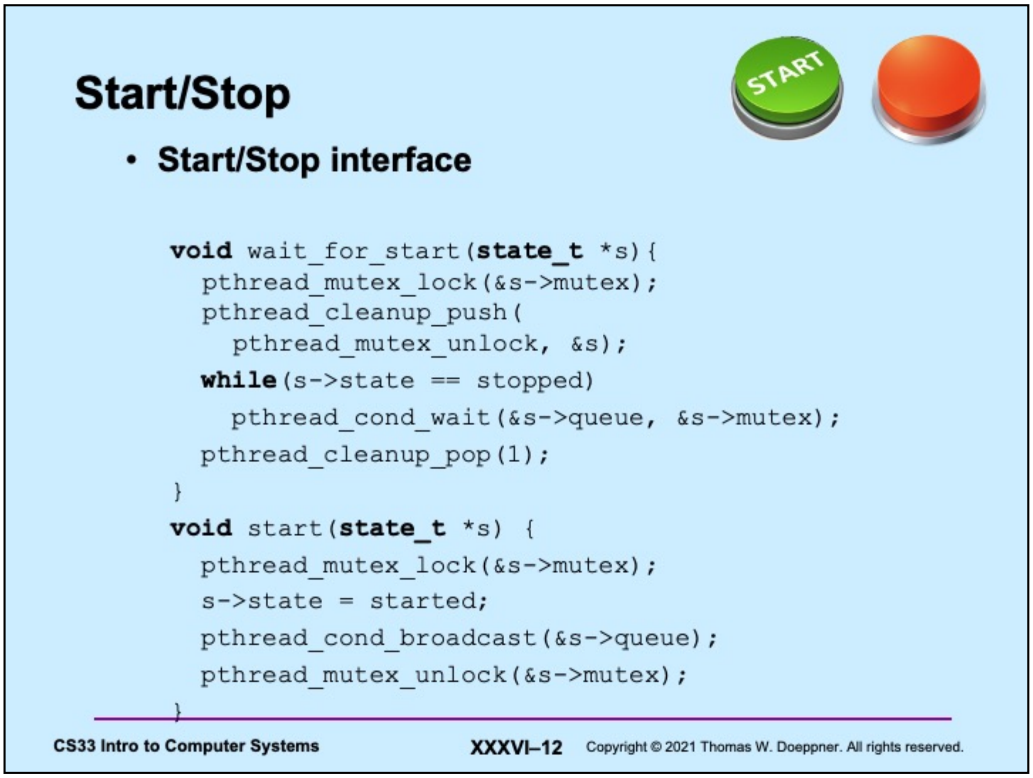
- **Start/Stop interface**

```
void wait_for_start(state_t *s){
  pthread_mutex_lock(&s->mutex);
  while(s->state == stopped)
    pthread_cond_wait(&s->queue,
      &s->mutex);
  pthread_mutex_unlock(&s->mutex);
}
void start(state_t *s) {
  pthread_mutex_lock(&s->mutex);
  s->state = started;
  pthread_cond_broadcast(&s->queue);
  pthread_mutex_unlock(&s->mutex);
}
```

Quiz 1

You're in charge of designing POSIX threads. Should *pthread_cond_wait* be a cancellation point?

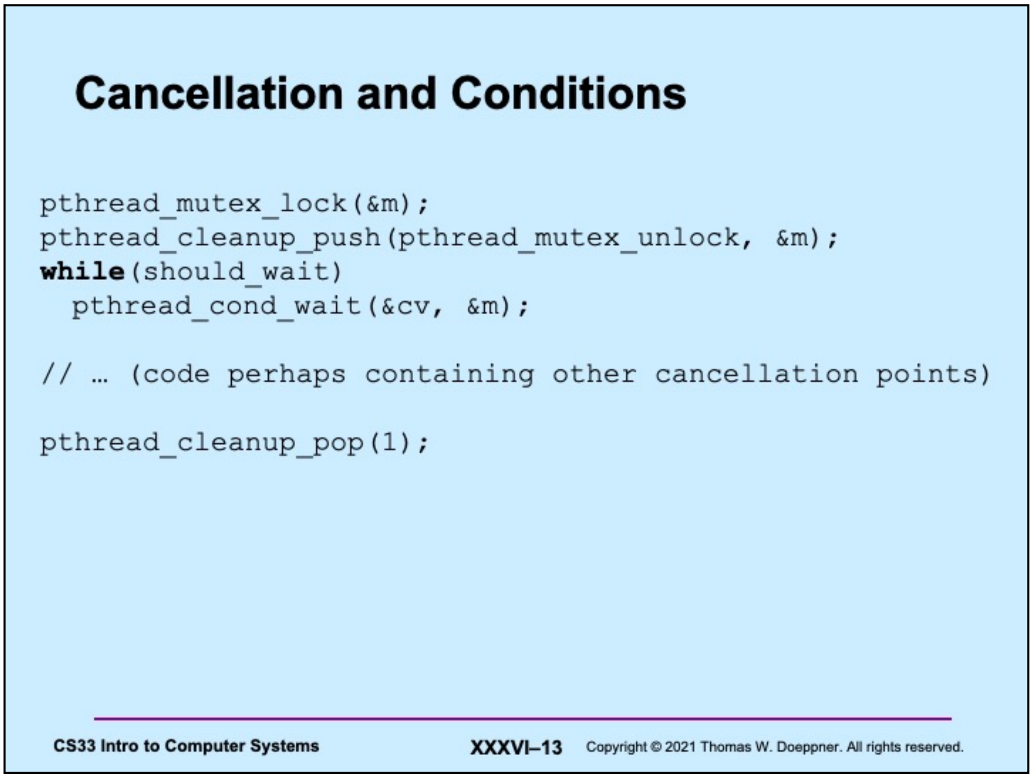
- **a) no**
- **b) yes; cancelled threads must acquire the mutex before invoking cleanup handler**
- **c) yes; but they don't acquire mutex**



Start/Stop

- **Start/Stop interface**

```
void wait_for_start(state_t *s){
  pthread_mutex_lock(&s->mutex);
  pthread_cleanup_push(
    pthread_mutex_unlock, &s);
  while(s->state == stopped)
    pthread_cond_wait(&s->queue, &s->mutex);
  pthread_cleanup_pop(1);
}
void start(state_t *s) {
  pthread_mutex_lock(&s->mutex);
  s->state = started;
  pthread_cond_broadcast(&s->queue);
  pthread_mutex_unlock(&s->mutex);
}
```

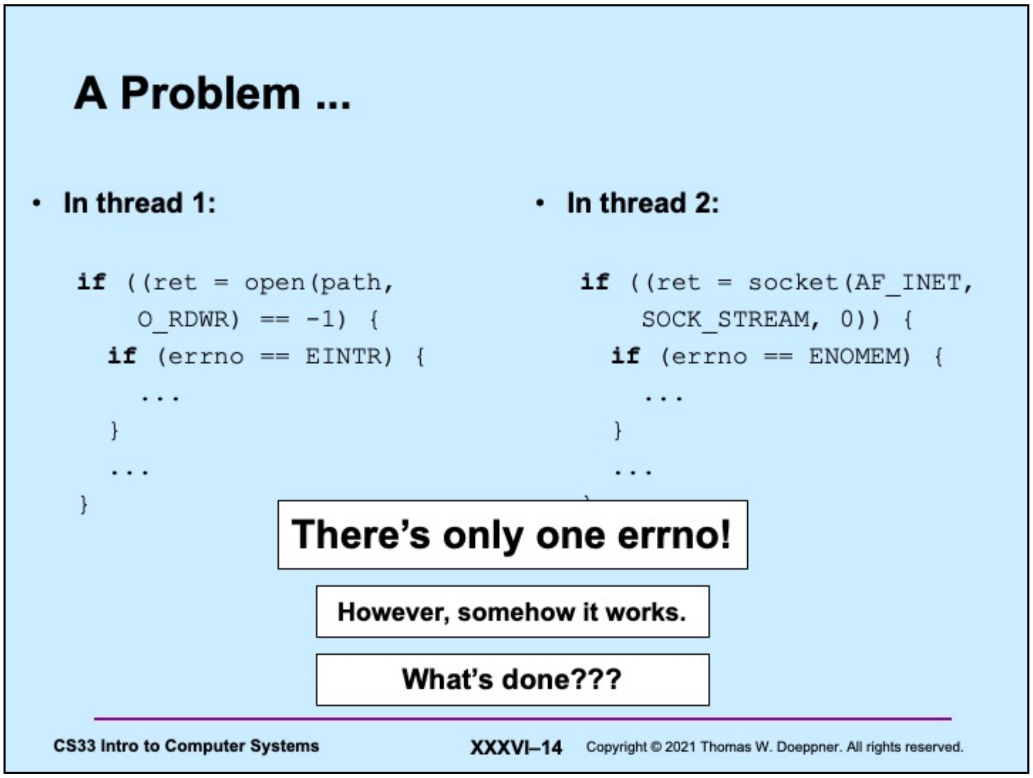


Cancellation and Conditions

```
pthread_mutex_lock(&m);
pthread_cleanup_push(pthread_mutex_unlock, &m);
while(should_wait)
  pthread_cond_wait(&cv, &m);

// … (code perhaps containing other cancellation points)

pthread_cleanup_pop(1);
```

In this example we handle cancels that might occur while a thread is blocked within **pthread_cond_wait**. Again, we assume the thread has cancels enabled and deferred. The thread first pushes a cleanup handler on its stack — in this case the cleanup handler unlocks the mutex. The thread then loops, calling **pthread_cond_wait**, a cancellation point. If it receives a cancel, the cleanup handler won't be called until the mutex has been reacquired. Thus, we are certain that when the cleanup handler is called, the mutex is locked.

What's important here is that we make sure the thread does not terminate without releasing its lock on the mutex *m*. If the thread acts on a cancel within **pthread_cond_wait** and the cleanup handler were invoked without first taking the mutex, this would be difficult to guarantee, since we wouldn't know if the thread had the mutex locked (and thus needs to unlock it) when it's in the cleanup handler.



A Problem ...

- In thread 1:

```
if ((ret = open(path,
    O_RDWR) == -1) {
  if (errno == EINTR) {
    ...
  }
  ...
}
```

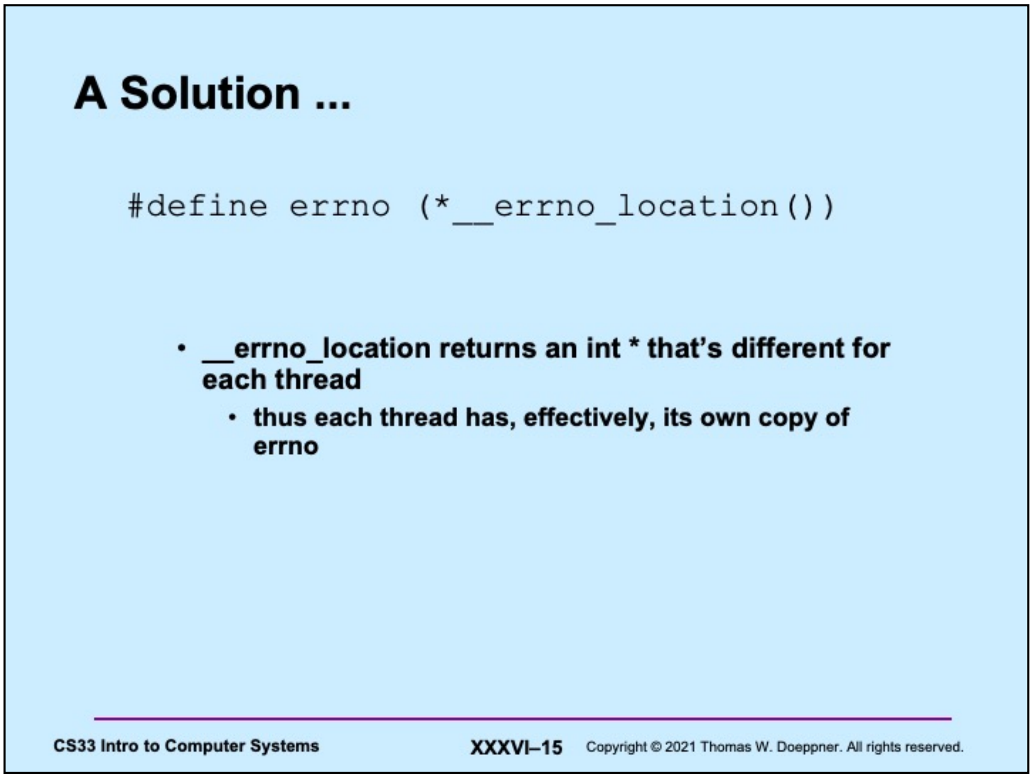
- In thread 2:

```
if ((ret = socket(AF_INET,
    SOCK_STREAM, 0)) {
  if (errno == ENOMEM) {
    ...
  }
  ...
}
```

There's only one errno!

However, somehow it works.

What's done???



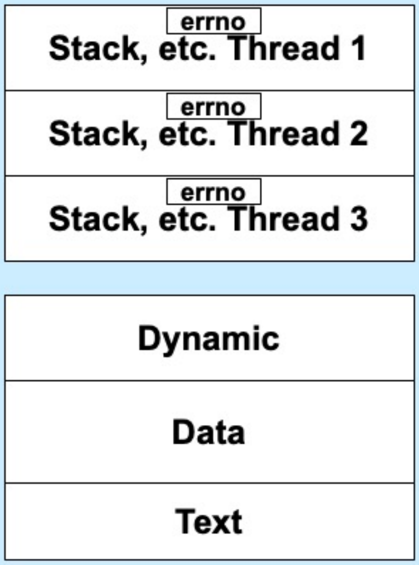
A Solution ...

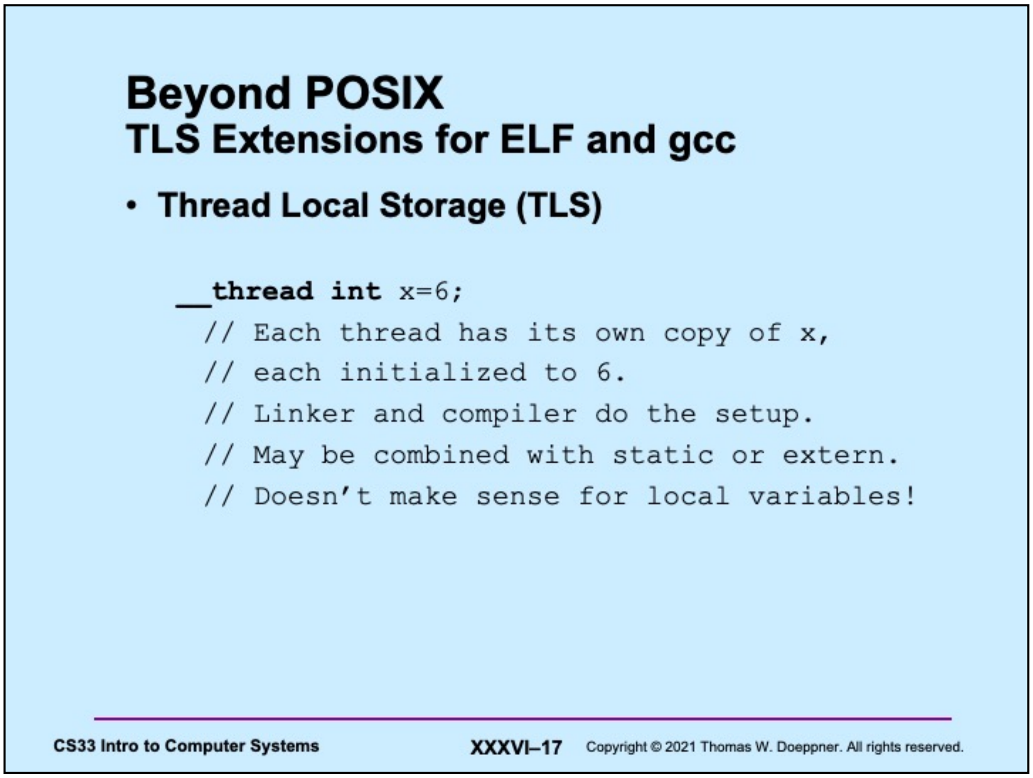
```
#define errno (*__errno_location())
```

- __errno_location returns an int * that's different for each thread
  - thus each thread has, effectively, its own copy of errno

When you give gcc the –pthread flag, it, among other things, defines some preprocessor variables that cause some code in the standard header files to be compiled (that otherwise wouldn't be). In particular the #define statement given in the slide is compiled.

Process Address Space



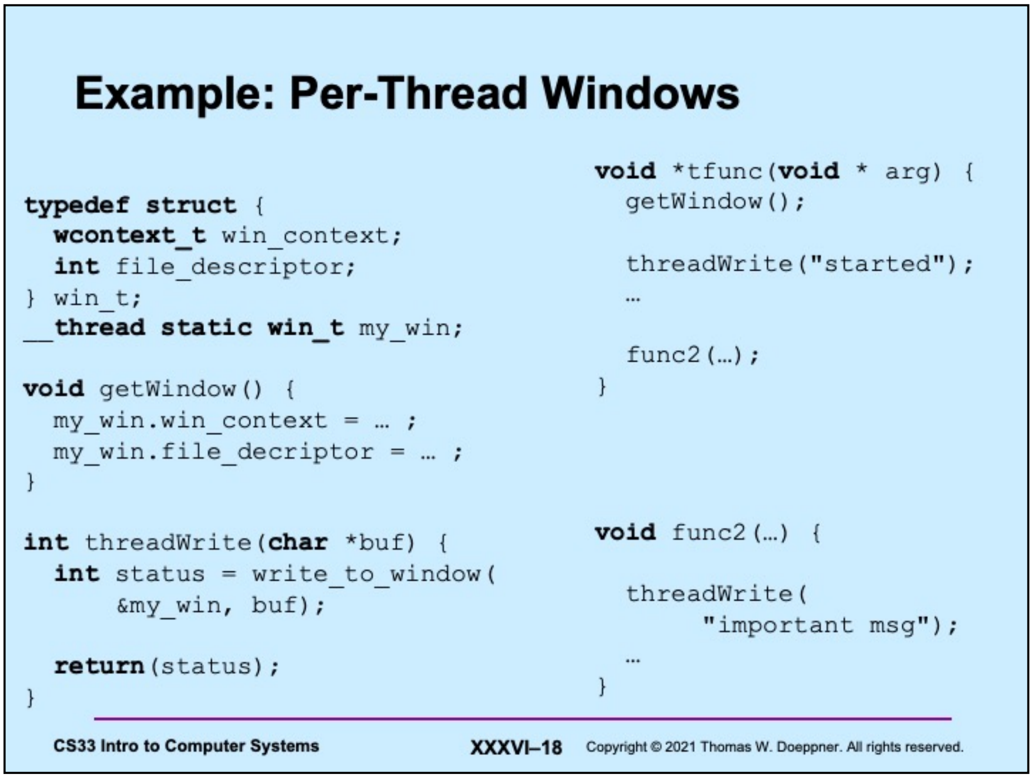


Beyond POSIX
TLS Extensions for ELF and gcc

- **Thread Local Storage (TLS)**

```
__thread int x=6;
  // Each thread has its own copy of x,
  // each initialized to 6.
  // Linker and compiler do the setup.
  // May be combined with static or extern.
  // Doesn’t make sense for local variables!
```

ELF stands for “executable and linking format” and is the standard format for executable and object files used on most Unix systems. The __**thread** attribute tells gcc that each thread is to have its own copy of the variable. A detailed description of how it is implemented can be found at http://people.redhat.com/drepper/tls.pdf.



Example: Per-Thread Windows

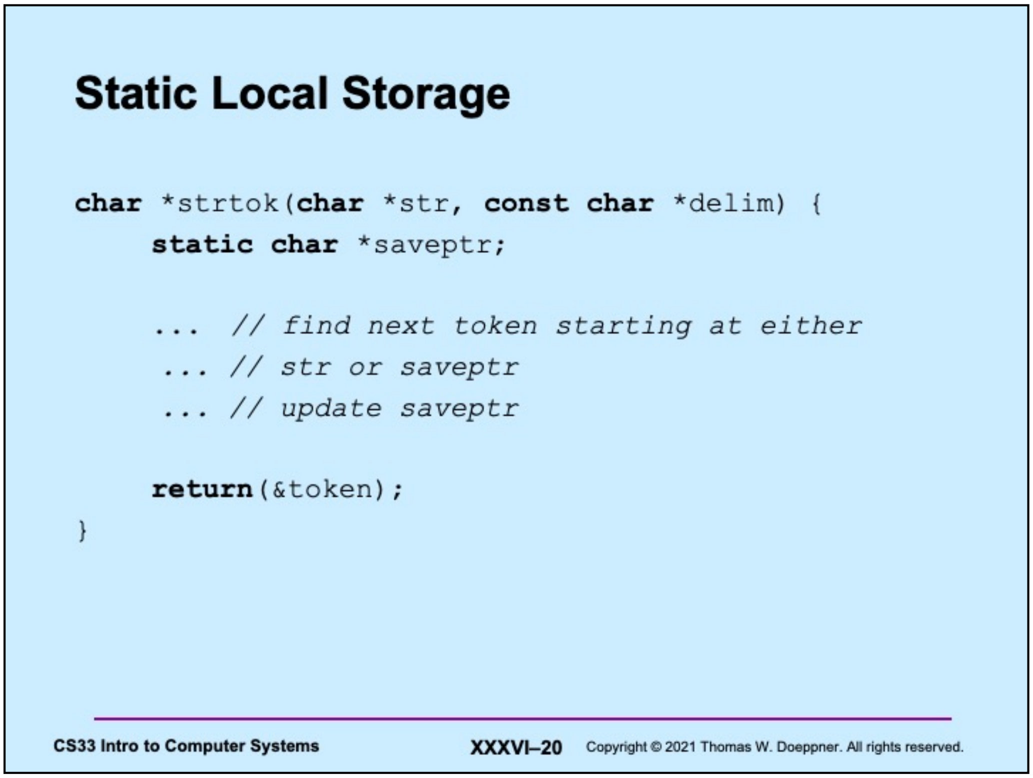
```
typedef struct {
  wcontext_t win_context;
  int file_descriptor;
} win_t;
__thread static win_t my_win;

void getWindow() {
  my_win.win_context = … ;
  my_win.file_decriptor = … ;
}

int threadWrite(char *buf) {
  int status = write_to_window(
      &my_win, buf);

  return(status);
}
```

```
void *tfunc(void * arg) {
  getWindow();

  threadWrite("started");
  …

  func2(…);
}


void func2(…) {

  threadWrite(
       "important msg");
  …
}
```

In this example, we put together per-thread windows for thread output. Threads call **getWindow** to set up a window for their exclusive use, then call **threadWrite** to send output to their windows. Individual threads can now set up their own windows and write to them without having to pass around information describing which are their windows.

Thread Safety

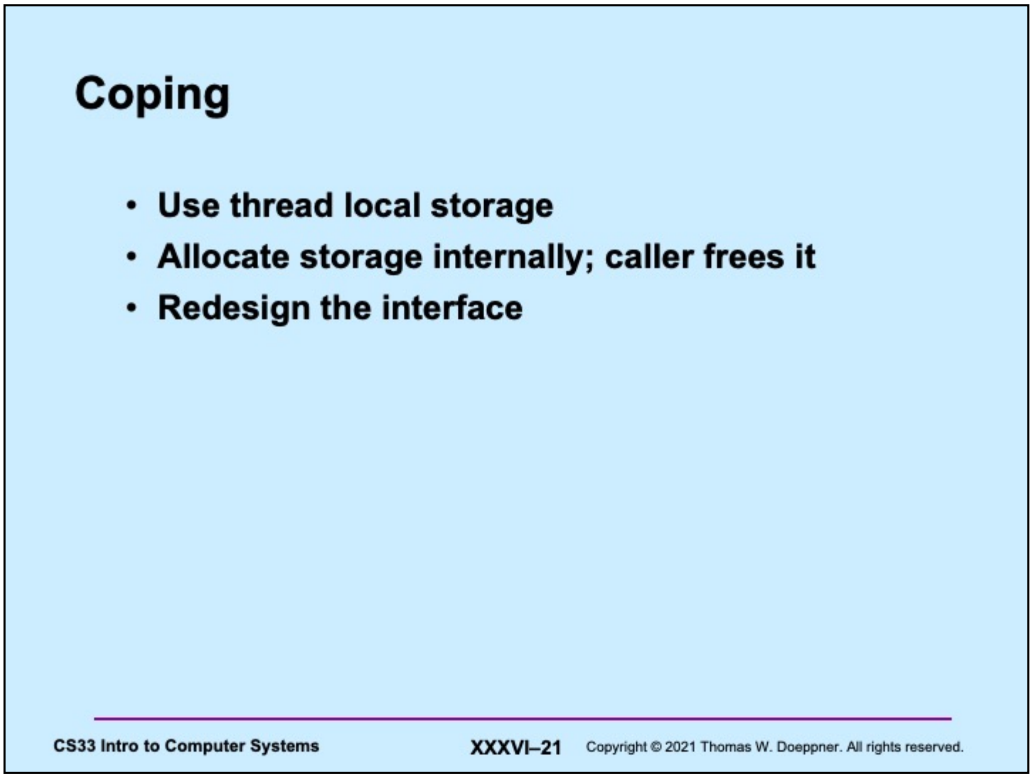
- Making the world safe for threads



Static Local Storage

```
char *strtok(char *str, const char *delim) {
    static char *saveptr;

    ...  // find next token starting at either
     ... // str or saveptr
     ... // update saveptr

    return(&token);
}
```

An example of the single-thread mentality in early Unix is the use of static local storage in a number of library routines. An example of this is **strtok**, which saves a pointer into the input string for use in future calls to the function (which we've used extensively in earlier projects). This works fine as long as just one thread is using the function, but fails if multiple threads use it – each will expect to find its own saved pointer in **saveptr**, but there's only one **saveptr**.

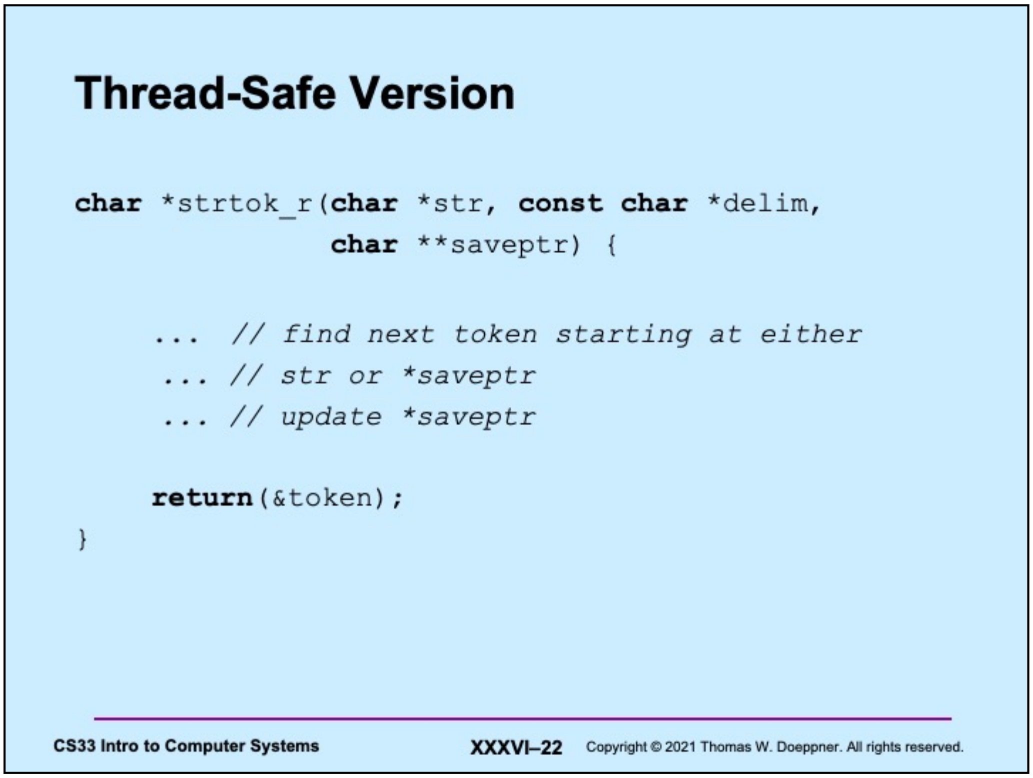


Coping

- Use thread local storage
- Allocate storage internally; caller frees it
- Redesign the interface

As the slide shows, there are at least three techniques for coping with this problem. We could use thread-local storage, but this would entail associating a fair amount of storage with each thread, even if it is not using **strtok**. We might simply allocate storage (via **malloc**) inside **strtok** and return a pointer to this storage. The problem with this is that the calls to **malloc** and **free** could turn out to be expensive. Furthermore, this makes it the caller's responsibility to free the storage, introducing a likely storage leak.

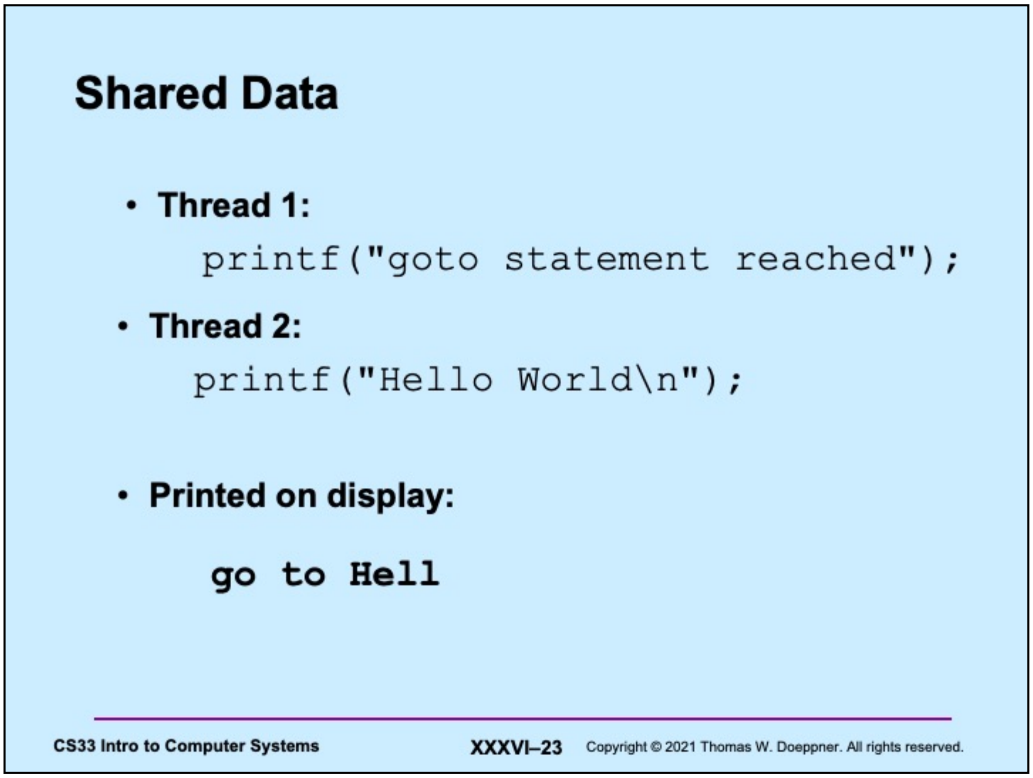
The solution taken is to redesign the interface. The "thread-safe" version is called **strtok_r** (the **r** stands for **reentrant**, an earlier term for "thread-safe"); it takes an additional parameter pointing to storage that holds saveptr. Thus the caller is responsible for both the allocation and the liberation of the storage containing saveptr; this storage is typically a local variable (allocated on the stack), so that its allocation and liberation overhead is negligible, at worst.



Thread-Safe Version

```
char *strtok_r(char *str, const char *delim,
               char **saveptr) {

    ...  // find next token starting at either
     ... // str or *saveptr
     ... // update *saveptr

    return(&token);
}
```

Here's the thread-safe version of **strtok**.



Shared Data

- **Thread 1:**

```
printf("goto statement reached");
```

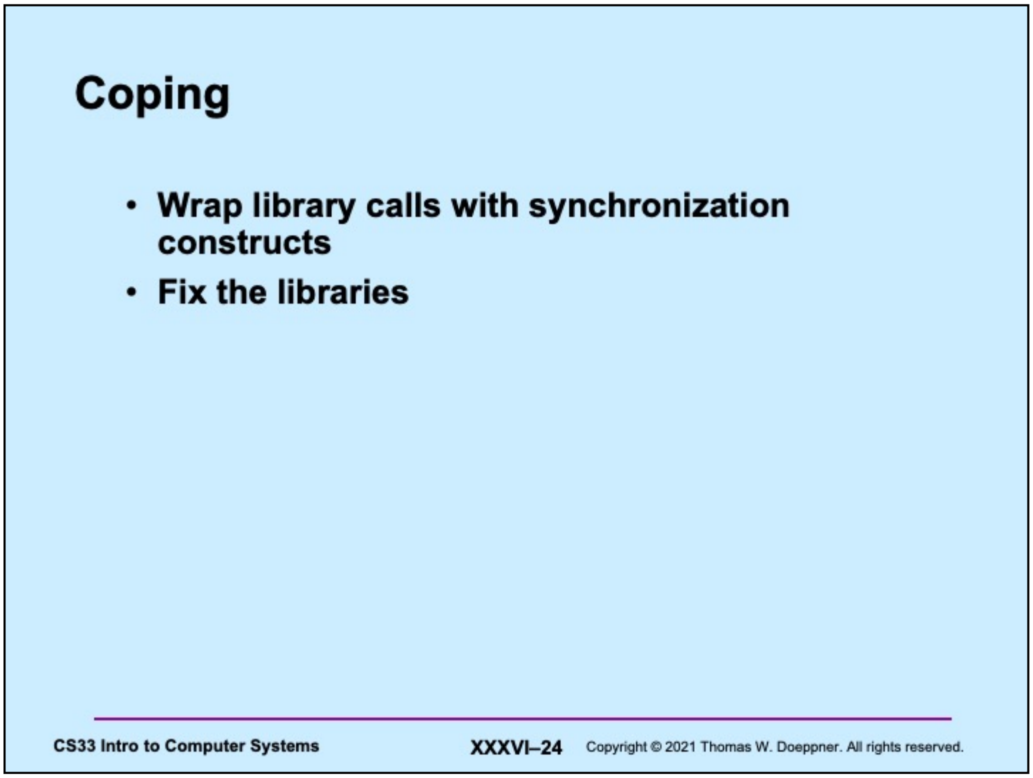
- **Thread 2:**

```
printf("Hello World\n");
```

- **Printed on display:**

```
go to Hell
```

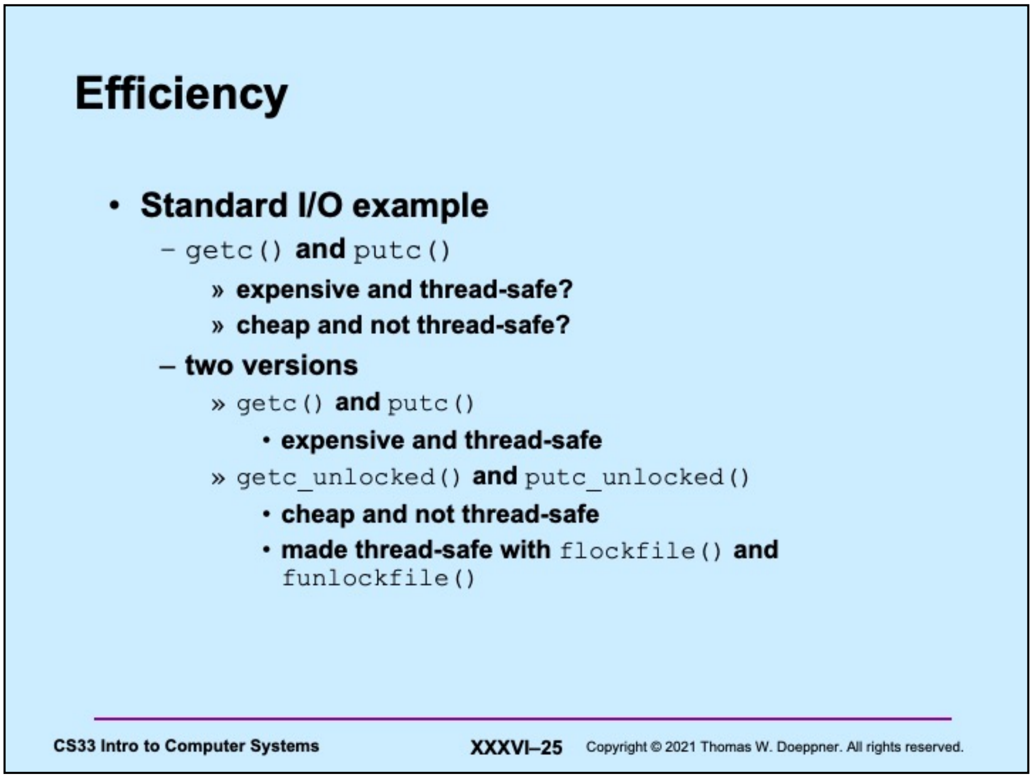
Yet another problem that arises when using libraries that were not designed for multithreaded programs concerns synchronization. The slide shows what might happen if one relied on the single-threaded versions of the standard I/O routines.



Coping

- Wrap library calls with synchronization constructs
- Fix the libraries

To deal with this **printf** problem, we must somehow add synchronization to **printf** (and all of the other standard I/O functions). A simple way to do this would be to supply wrappers for all of the standard I/O functions ensuring that only one thread is operating on any particular stream at a time. A better way would be to do the same sort of thing by fixing the functions themselves, rather than supplying wrappers (this is what is done in most implementations).

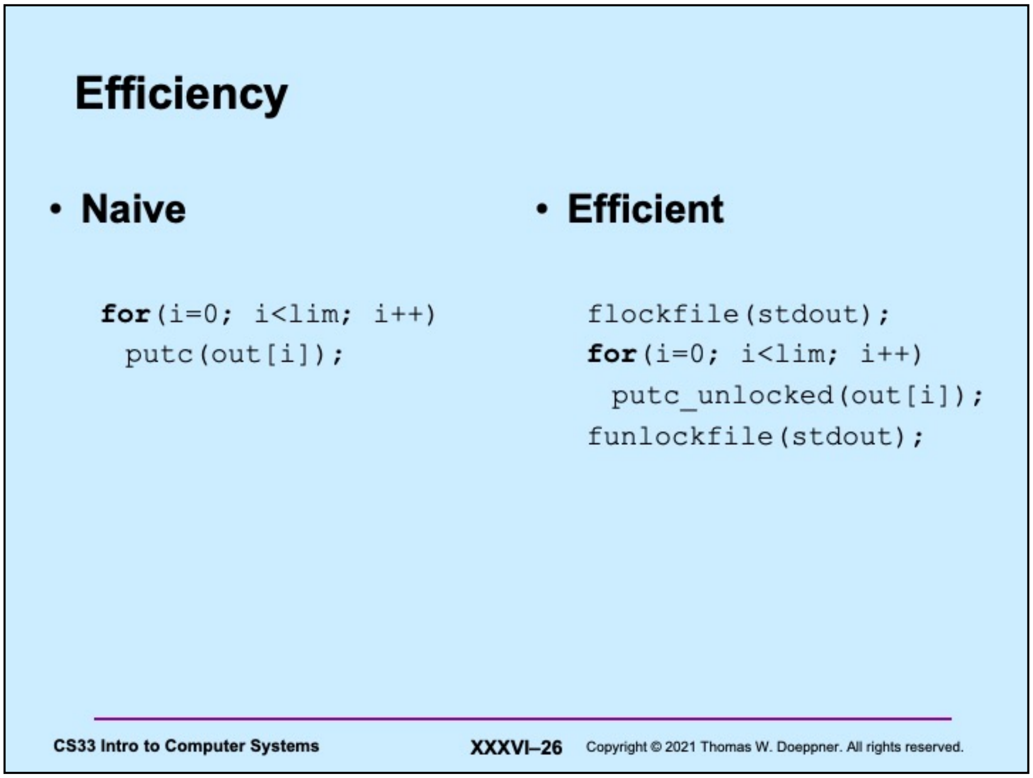


Efficiency

- **Standard I/O example**
  - `getc()` **and** `putc()`
    - » **expensive and thread-safe?**
    - » **cheap and not thread-safe?**
  - **two versions**
    - » `getc()` **and** `putc()`
      - **expensive and thread-safe**
    - » `getc_unlocked()` **and** `putc_unlocked()`
      - **cheap and not thread-safe**
      - **made thread-safe with** `flockfile()` **and** `funlockfile()`

After making a library thread-safe, we may discover that many functions have become too slow. For example, the standard-I/O functions **getc** and **putc** are expected to be fast — they are usually implemented as macros. But once we add the necessary synchronization, they become rather sluggish — much too slow to put in our innermost loops. However, if we are aware of and willing to cope with the synchronization requirements ourselves, we can produce code that is almost as efficient as the single-threaded code without synchronization requirements.

The POSIX-threads specification includes unsynchronized versions of **getc** and **putc** — **getc_unlocked** and **putc_unlocked**. These are exactly the same code as the single-threaded **getc** and **putc**. To use these new functions, one must take care to handle the synchronization oneself. This is accomplished with **flockfile** and **funlockfile**.



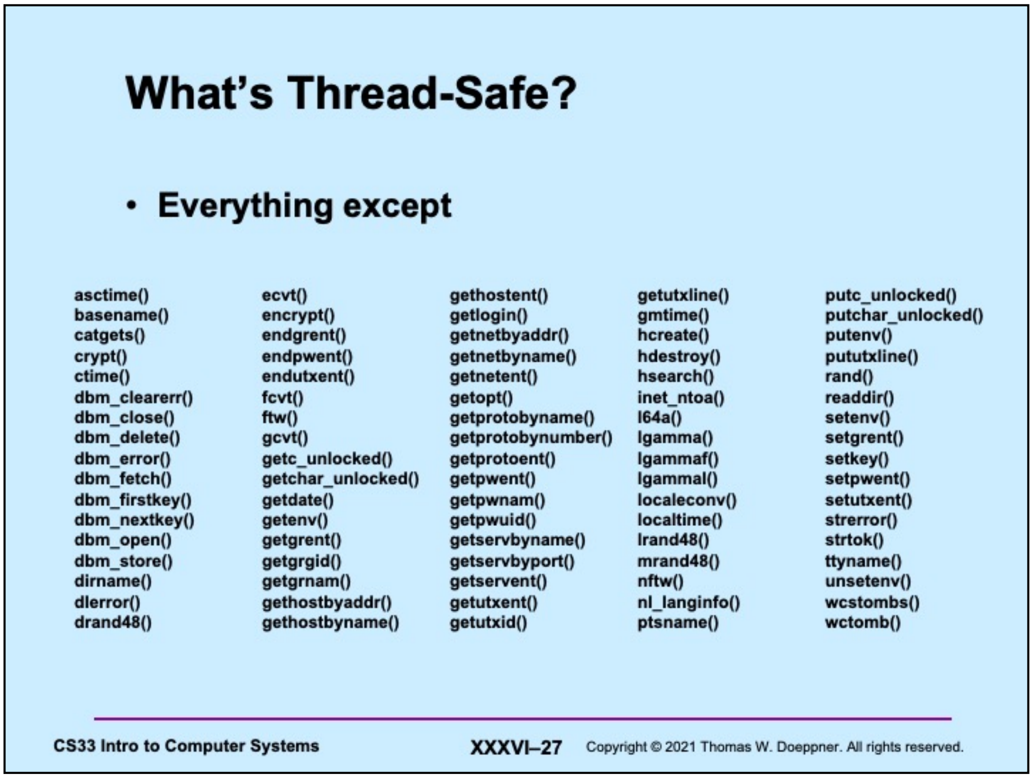
Efficiency

- **Naive**

```
for(i=0; i<lim; i++)
 putc(out[i]);
```

- **Efficient**

```
flockfile(stdout);
for(i=0; i<lim; i++)
 putc_unlocked(out[i]);
funlockfile(stdout);
```



What's Thread-Safe?

- Everything except

asctime()
basename()
catgets()
crypt()
ctime()
dbm_clearerr()
dbm_close()
dbm_delete()
dbm_error()
dbm_fetch()
dbm_firstkey()
dbm_nextkey()
dbm_open()
dbm_store()
dirname()
dlerror()
drand48()
ecvt()
encrypt()
endgrent()
endpwent()
endutxent()
fcvt()
ftw()
gcvt()
getc_unlocked()
getchar_unlocked()
getdate()
getenv()
getgrent()
getgrgid()
getgrnam()
gethostbyaddr()
gethostbyname()
gethostent()
getlogin()
getnetbyaddr()
getnetbyname()
getnetent()
getopt()
getprotobyname()
getprotobynumber()
getprotoent()
getpwent()
getpwnam()
getpwuid()
getservbyname()
getservbyport()
getservent()
getutxent()
getutxid()
getutxline()
gmtime()
hcreate()
hdestroy()
hsearch()
inet_ntoa()
l64a()
lgamma()
lgammaf()
lgammal()
localeconv()
localtime()
lrand48()
mrand48()
nftw()
nl_langinfo()
ptsname()
putc_unlocked()
putchar_unlocked()
putenv()
pututxline()
rand()
readdir()
setenv()
setgrent()
setkey()
setpwent()
setutxent()
strerror()
strtok()
ttyname()
unsetenv()
wcstombs()
wctomb()

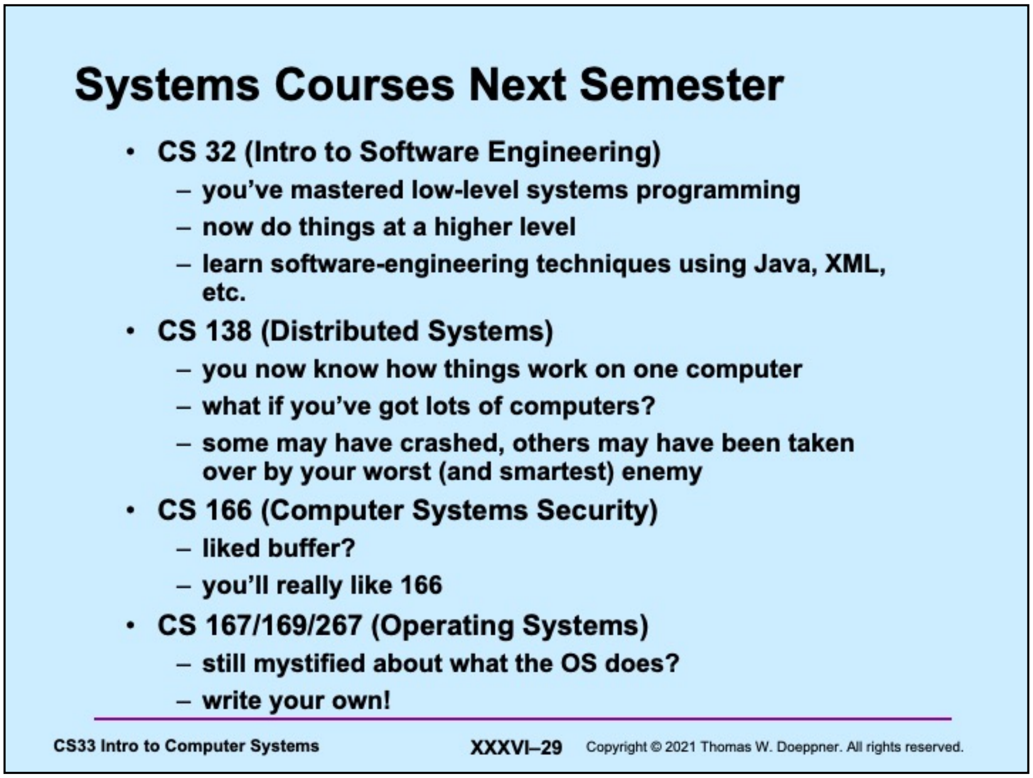
According to IEEE Std. 1003.1 (POSIX), all functions it specifies must be thread-safe, except for those listed above.

You'll Soon Finish CS 33 …

- You might
  - celebrate

  - take another systems course
    - » 32
    - » 138
    - » 166
    - » 167
    - » 168

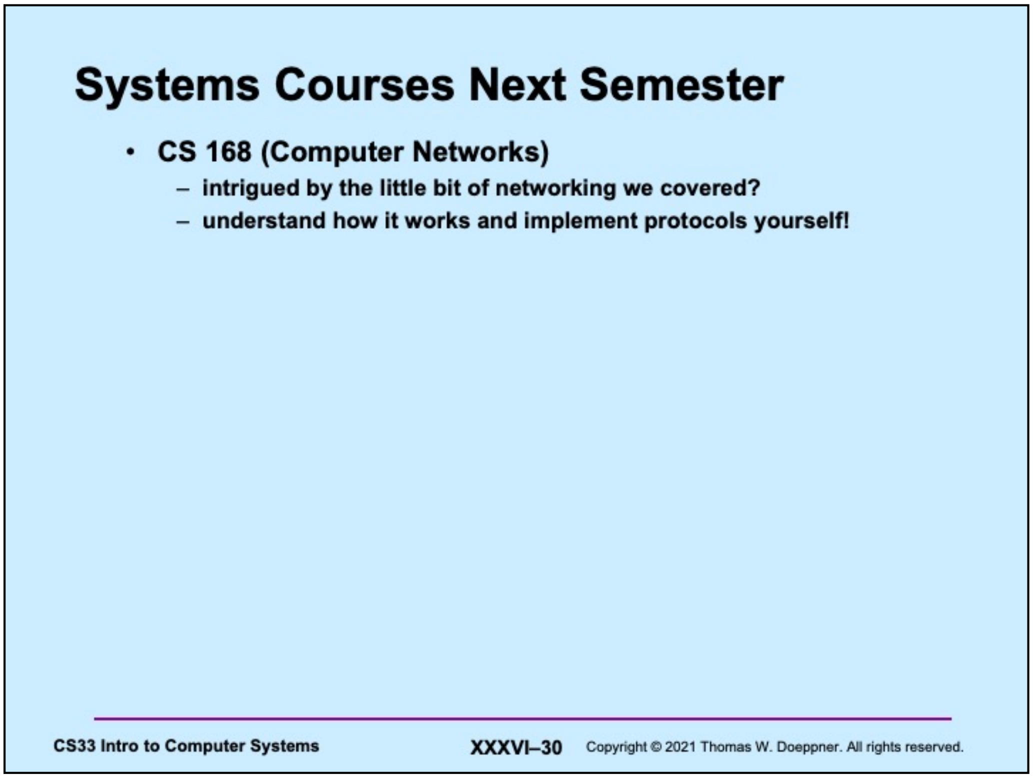
  - become a 33 TA



Systems Courses Next Semester

- CS 32 (Intro to Software Engineering)
  - you've mastered low-level systems programming
  - now do things at a higher level
  - learn software-engineering techniques using Java, XML, etc.
- CS 138 (Distributed Systems)
  - you now know how things work on one computer
  - what if you've got lots of computers?
  - some may have crashed, others may have been taken over by your worst (and smartest) enemy
- CS 166 (Computer Systems Security)
  - liked buffer?
  - you'll really like 166
- CS 167/169/267 (Operating Systems)
  - still mystified about what the OS does?
  - write your own!

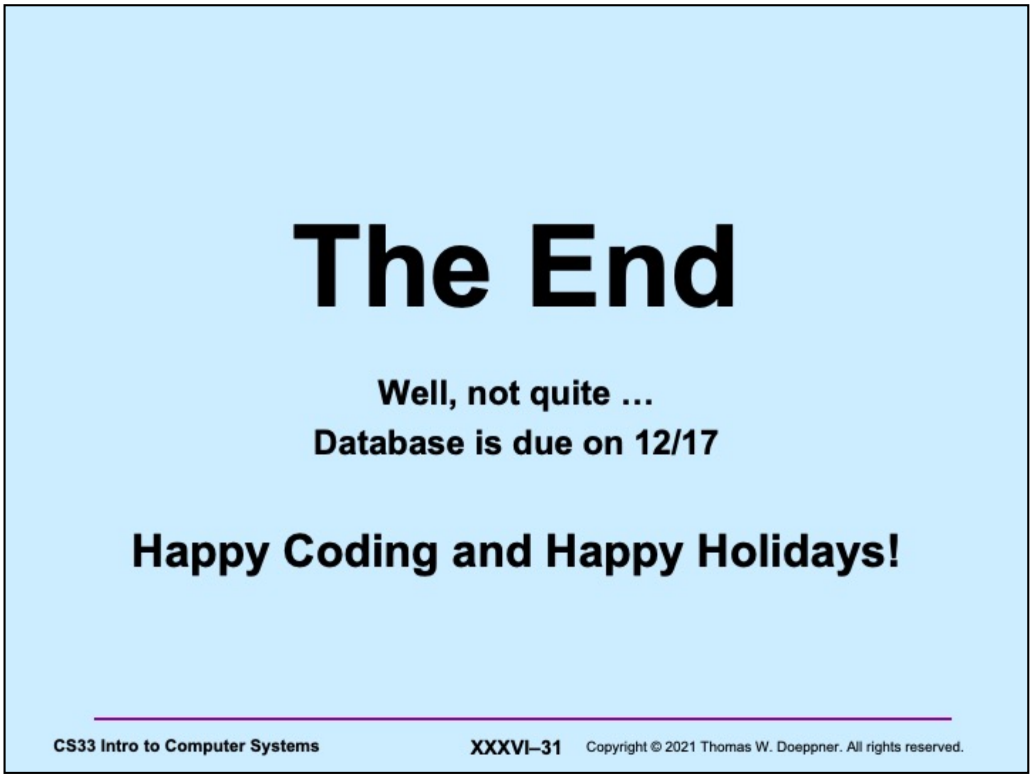
267 is for graduate students only and combines 167 and 169.



Systems Courses Next Semester

- CS 168 (Computer Networks)
  - intrigued by the little bit of networking we covered?
  - understand how it works and implement protocols yourself!

We don't know yet whether 168 will be offered in the next academic year.



The End

Well, not quite …

Database is due on 12/17

Happy Coding and Happy Holidays!