EE309 Advanced Programming Techniques for EE

Lecture 14: Concurrent Programming INSU YUN (윤인수)

School of Electrical Engineering, KAIST

[Slides from 15-213: Introduction to Computer Systems at CMU]

Concurrent Programming is Hard!

- The human mind tends to be sequential
- The notion of time is often misleading
- Thinking about all possible sequences of events in a computer system is at least error prone and frequently impossible

Concurrent Programming is Hard!

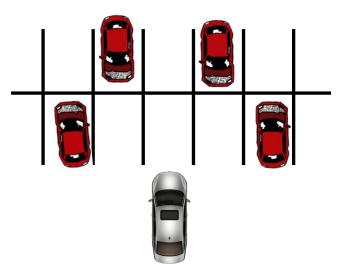
Classical problem classes of concurrent programs:

- Races: outcome depends on arbitrary scheduling decisions elsewhere in the system
 - Example: who gets the last seat on the airplane?
- **Deadlock:** improper resource allocation prevents forward progress
 - Example: traffic gridlock
- Livelock / Starvation / Fairness: external events and/or system scheduling decisions can prevent sub-task progress
 - Example: people always jump in front of you in line

Data Race







KAIST

Deadlock





Deadlock

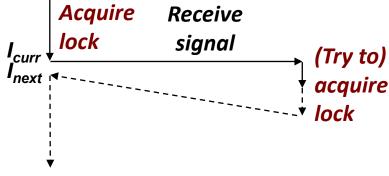
- Example from signal handlers.
- Why don't we use printf in handlers?



```
void catch_child(int signo) {
    printf("Child exited!\n"); // this call may reenter printf/puts! BAD! DEADLOCK!
    while (waitpid(-1, NULL, WNOHANG) > 0) continue; // reap all children
}
```

Printf code:

- Acquire lock
- Do something
- Release lock



Deadlock

- Example from signal handlers.
- Why don't we use printf in handlers?



(Try to)

acquire

lock

Deadlocked!

```
void catch_child(int signo) {
    printf("Child exited!\n"); // this call may reenter printf/puts! BAD! DEADLOCK!
    while (waitpid(-1, NULL, WNOHANG) > 0) continue; // reap all children
}
Acquire Receive
```

signal

lock

```
Printf code:
```

- Acquire lock
- Do something
- Release lock
- What if signal handler interrupts call to printf?

Testing Printf Deadlock

```
static void sigchld(int unused) {
    int status;
   pid t pid;
    while ((pid = waitpid(-1, &status, WNOHANG)) > 0) {
        printf("Child %d exited with status %04x\n", pid, status);
    }
}
int main(void) {
    signal(SIGCHLD, sigchld);
    for (int i = 0; i < 1000000; i++) {</pre>
        pid t pid = fork();
        if (pid == 0)
            exit(0);
        // in parent
        printf("Child #%d=%d started\n",
               i, pid);
    }
    return 0;
}
```

Testing Printf Deadlock

```
static void sigchld(int unused) {
    int status;
   pid t pid;
    while ((pid = waitpid(-1, \&status, WNOHANG)) > 0) {
        printf("Child %d exited with status %04x\n", pid, status);
    }
}
                                 Child #0=1234 started
int main(void) {
    signal(SIGCHLD, sigchld);
                                 Child #1=1235 started
    for (int i = 0; i < 1000000; Child #2=1236 started
        pid t pid = fork();
                                 Child #3=1237 started
        if (pid == 0)
                                 Child 1234 exited with status 0000
            exit(0);
                                 Child #4=1238 started
        // in parent
                                 Child 1235 exited with status 0000
        printf("Child #%d=%d star Child 1236 exited with status 0000
               i, pid);
    return 0;
                                  Child #3566=16979 started
                                  and then, silence
```

Testing Printf Deadlock

```
static void sigchld(int unused) {
    int status;
    pid_t pid;
    while ((pid = waitpid(-1, &status, WNOHANG)) > 0) {
        printf("Child %d exited with status %04x\n", pid, status);
    }
}
int main(void) {
    (gdb) bt
    #0 0x00007ffff7b197fc in __lll_lock_wait_private ()
    #1 0x00007ffff7a5b00e in _L_lock_1177 ()
```

```
#2 0x00007ffff7a557f4 in _IO_vfprintf_internal ()
```

```
#3 0 \times 00007 ffff7a604e9 in printf (
```

```
format="Child %d exited with status %04x\n")
```

```
#4 0x000000000000678 in sigchld ()
```

```
#5 <signal handler called>
```

}

```
#6 0x00007ffff7a5583f in _IO_vfprintf_internal ()
```

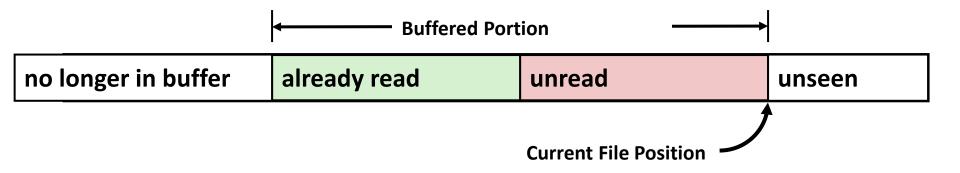
```
#7 0x00007ffff7a604e9 in printf (
```

```
format="Child #%d=%d started\n")
```

```
#8 0x00000000004006d2 in main ()
```

Why Does Printf require Locks?

Printf (and fprintf, sprintf) implement buffered I/O



Require locks to access the shared buffers

KAIST

Livelock

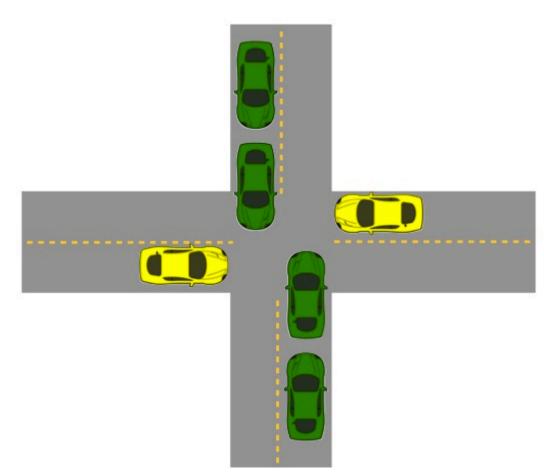


KAIST

Livelock



Starvation



- Yellow must yield to green
- Continuous stream of green cars
- Overall system makes progress, but some individuals wait indefinitely

Concurrent Programming is Hard!

Classical problem classes of concurrent programs:

- **Races:** outcome depends on arbitrary scheduling decisions elsewhere in the system
 - Example: who gets the last seat on the airplane?
- **Deadlock:** improper resource allocation prevents forward progress
 - Example: traffic gridlock
- Livelock / Starvation / Fairness: external events and/or system scheduling decisions can prevent sub-task progress
 - Example: people always jump in front of you in line

Many aspects of concurrent programming are beyond the scope of our course..

- but, not all ③
- We'll cover some of these as

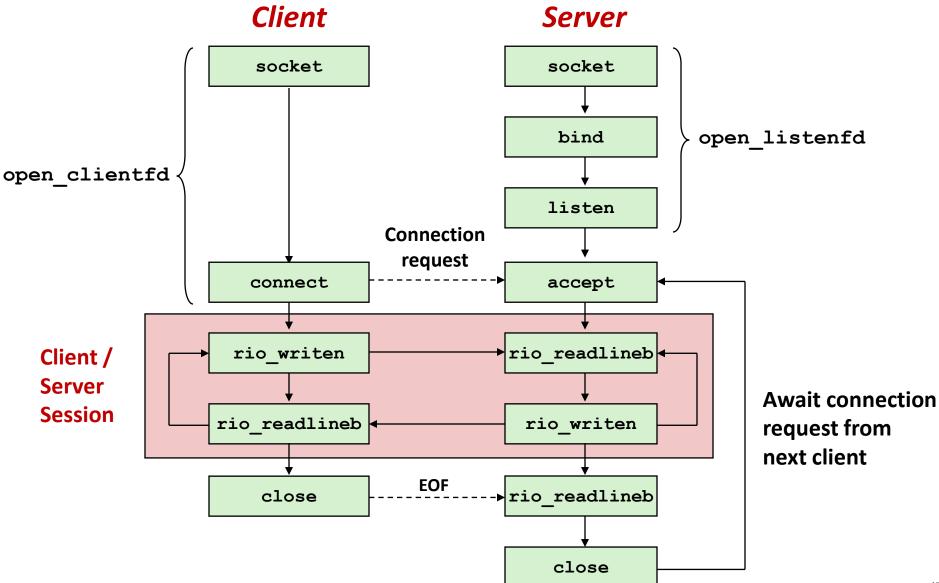
Do Activity 1 and 2: All Problems

Concurrent Programming is Hard!

It may be hard, but ...

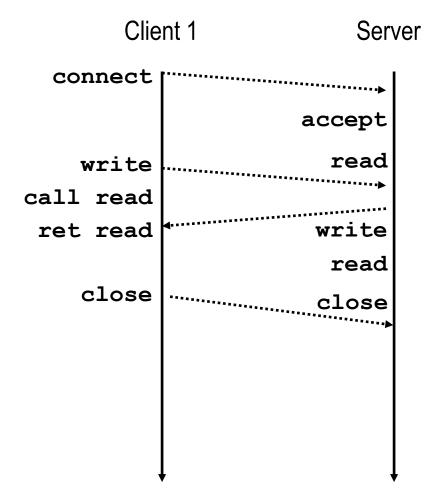
it can be useful and more and more necessary!

Reminder: Iterative Echo Server



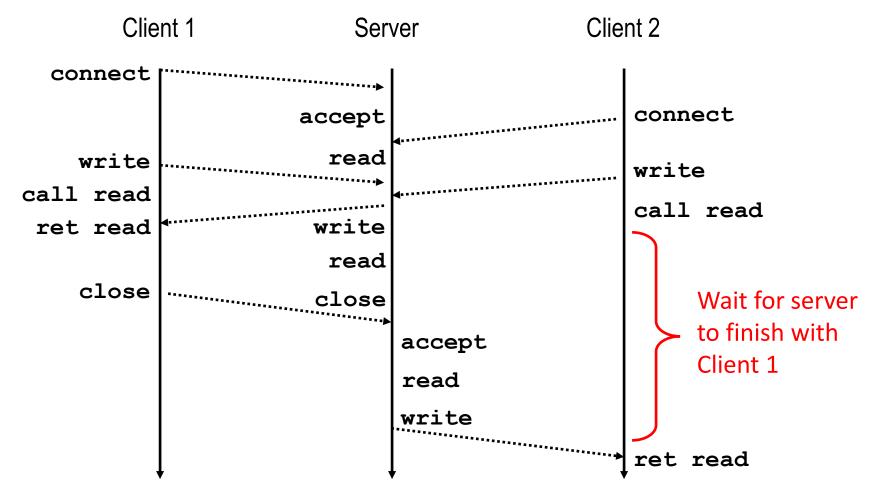
Iterative Servers

Iterative servers process one request at a time



Iterative Servers

Iterative servers process one request at a time



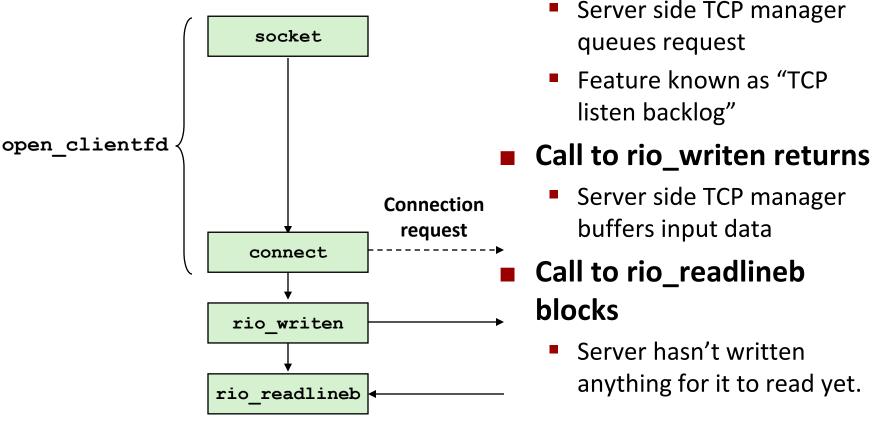
Call to connect returns

yet accepted

Even though connection not

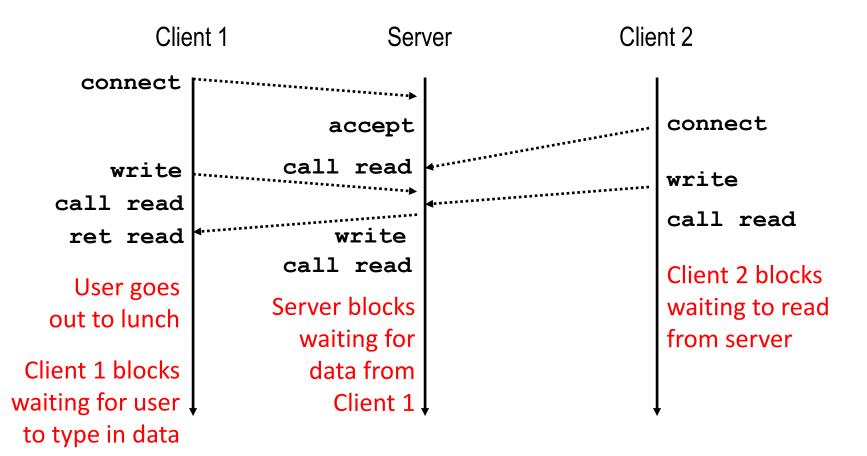
Where Does Second Client Block?

Second client attempts to connect to iterative server



Client

Fundamental Flaw of Iterative Servers



Solution: use concurrent servers instead

 Concurrent servers use multiple concurrent flows to serve multiple clients at the same time

Approaches for Writing Concurrent Servers

Allow server to handle multiple clients concurrently

1. Process-based

- Kernel automatically interleaves multiple logical flows
- Each flow has its own private address space

2. Event-based

- Programmer manually interleaves multiple logical flows
- All flows share the same address space
- Uses technique called I/O multiplexing

3. Thread-based

- Kernel automatically interleaves multiple logical flows
- Each flow shares the same address space
- Hybrid of of process-based and event-based

Approaches for Writing Concurrent Servers

Allow server to handle multiple clients concurrently

1. Process-based

- Kernel automatically interleaves multiple logical flows
- Each flow has its own private address space

2. Event-based

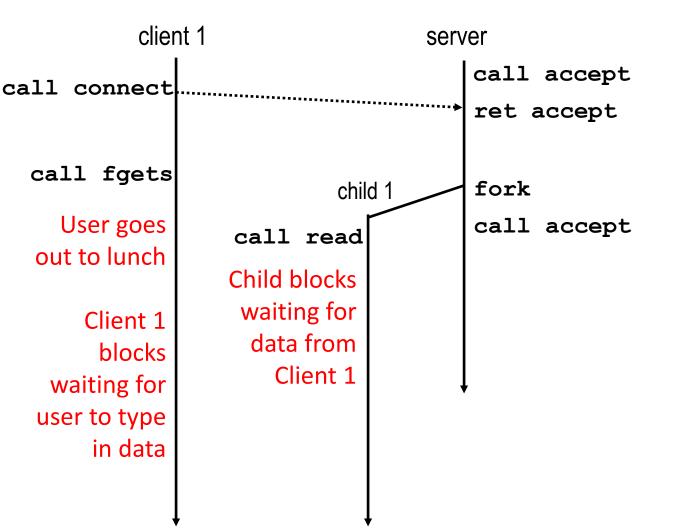
- Programmer manually interleaves multiple logical flows
- All flows share the same address space
- Uses technique called I/O multiplexing

3. Thread-based

- Kernel automatically interleaves multiple logical flows
- Each flow shares the same address space
- Hybrid of of process-based and event-based

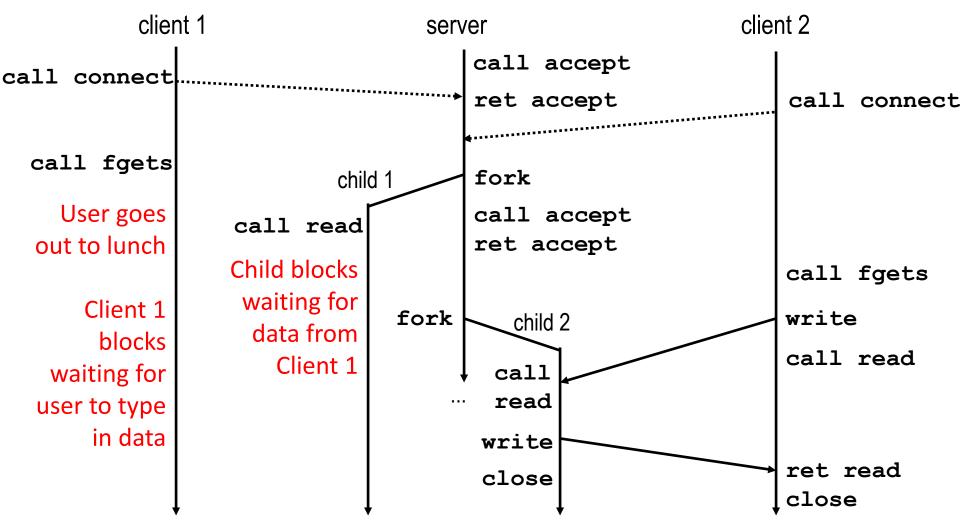
Approach #1: Process-based Servers

Spawn separate process for each client



Approach #1: Process-based Servers

Spawn separate process for each client



Iterative Echo Server

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
        echo(connfd);
        Close (connfd);
     }
     exit(0);
                Accept a connection request
                Handle echo requests until client terminates
                                                                echoserverp.c
```

}

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
            echo(connfd); /* Child services client */
            Close(connfd); /* child closes connection with client */
            exit(0);
```

echoserverp.c

}

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
        if (Fork() == 0) {
            echo(connfd); /* Child services client */
            Close (connfd); /* Child closes connection with client */
                            /* Child exits */
            exit(0);
        }
```

echoserverp.c

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
        if (Fork() == 0) {
            echo(connfd); /* Child services client */
            Close(connfd); /* Child closes connection with client */
                            /* Child exits */
            exit(0);
        }
        Close(connfd); /* Parent closes connected socket (important!) */
    }
}
                                                               echoserverp.c
                   Whv?
```

}

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
        if (Fork() == 0) {
            Close (listenfd); /* Child closes its listening socket */
            echo(connfd); /* Child services client */
            Close (connfd); /* Child closes connection with client */
            exit(0); /* Child exits */
        }
        Close(connfd); /* Parent closes connected socket (important!) */
    }
```

Process-Based Concurrent Echo Server

}

```
int main(int argc, char **argv)
{
    int listenfd, connfd;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
    Signal(SIGCHLD, sigchld handler);
    listenfd = Open listenfd(argv[1]);
    while (1) {
        clientlen = sizeof(struct sockaddr storage);
        connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
        if (Fork() == 0) {
            Close (listenfd); /* Child closes its listening socket */
            echo(connfd); /* Child services client */
            Close (connfd); /* Child closes connection with client */
            exit(0); /* Child exits */
        }
        Close(connfd); /* Parent closes connected socket (important!) */
    }
```

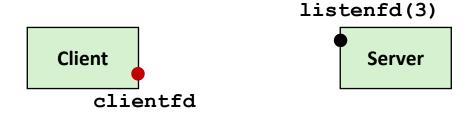
echoserverp.c

Process-Based Concurrent Echo Server (cont)

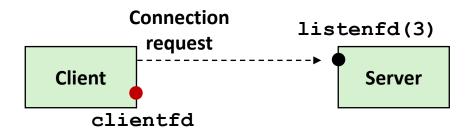
```
void sigchld_handler(int sig)
{
    while (waitpid(-1, 0, WNOHANG) > 0)
        ;
        return;
} echoserverp.c
```

Reap all zombie children

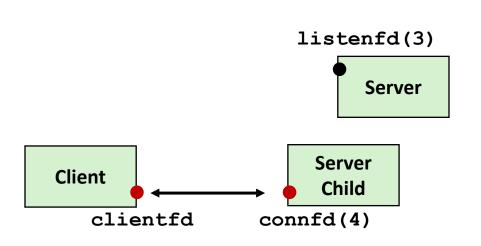
Concurrent Server: accept Illustrated



1. Server blocks in accept, waiting for connection request on listening descriptor listenfd

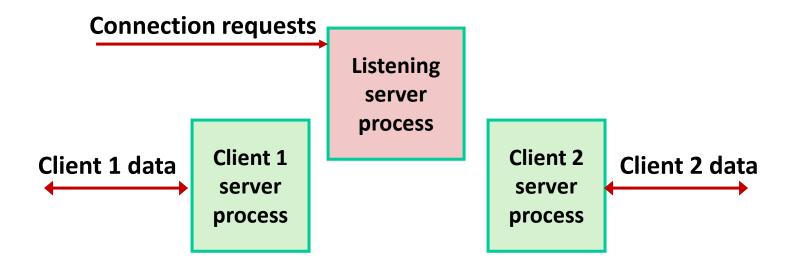


2. Client makes connection request by calling connect



3. Server returns connfd from accept. Forks child to handle client. Connection is now established between clientfd and connfd

Process-based Server Execution Model



- Each client handled by independent child process
- No shared state between them
- Both parent & child have copies of listenfd and connfd
 - Parent must close connfd
 - Child should close listenfd

Issues with Process-based Servers

- Listening server process must reap zombie children
 - to avoid fatal memory leak

Parent process must close its copy of connfd

- Kernel keeps reference count for each socket/open file
- After fork, refcnt(connfd) = 2
- Connection will not be closed until refcnt (connfd) = 0

Pros and Cons of Process-based Servers

+ Handle multiple connections concurrently

- + Clean sharing model
 - descriptors (no)
 - file tables (yes)
 - global variables (no)
- + Simple and straightforward
- Additional overhead for process control
- Nontrivial to share data between processes
 - (This example too simple to demonstrate)

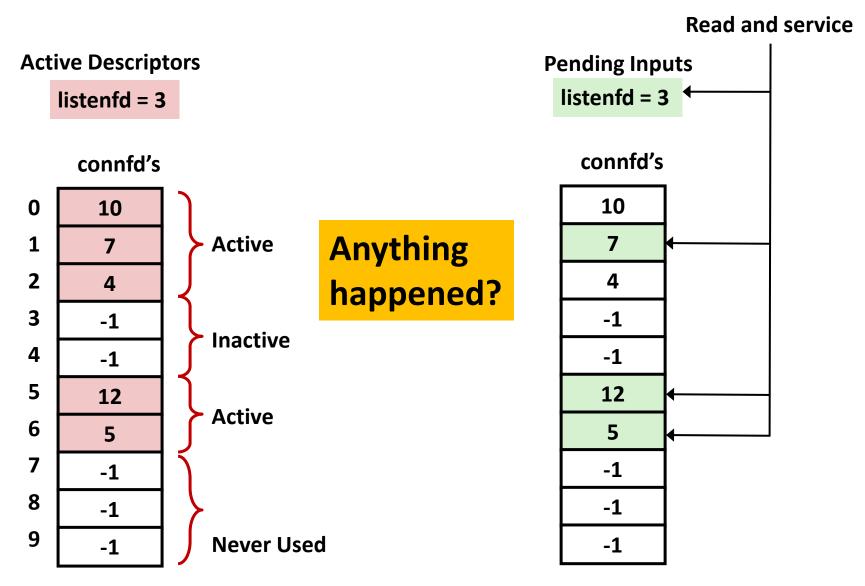
Approach #2: Event-based Servers

Server maintains set of active connections

- Array of connfd's
- Repeat:
 - Determine which descriptors (connfd's or listenfd) have pending inputs
 - e.g., using select function
 - arrival of pending input is an *event*
 - If listenfd has input, then accept connection
 - and add new connfd to array
 - Service all connfd's with pending inputs

Details for select-based server in book

I/O Multiplexed Event Processing



Pros and Cons of Event-based Servers

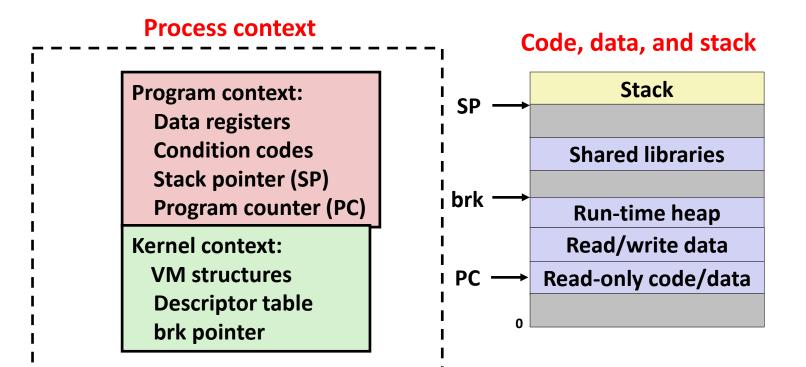
- + One logical control flow and address space.
- + Can single-step with a debugger.
- + No process or thread control overhead.
 - Design of choice for high-performance Web servers and search engines.
 e.g., Node.js, nginx, Tornado
- Significantly more complex to code than process- or threadbased designs.
- Hard to provide fine-grained concurrency
 - E.g., how to deal with partial HTTP request headers
- Cannot take advantage of multi-core
 - Single thread of control

Approach #3: Thread-based Servers

- Very similar to approach #1 (process-based)
 - ...but using threads instead of processes

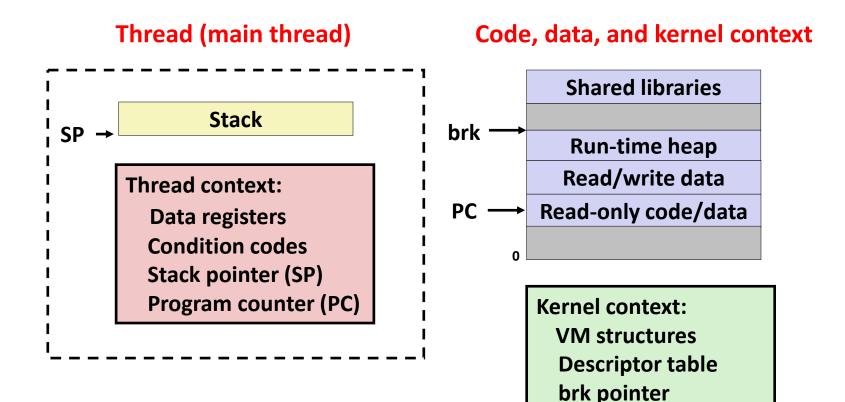
Traditional View of a Process

Process = process context + code, data, and stack



Alternate View of a Process

Process = thread + code, data, and kernel context



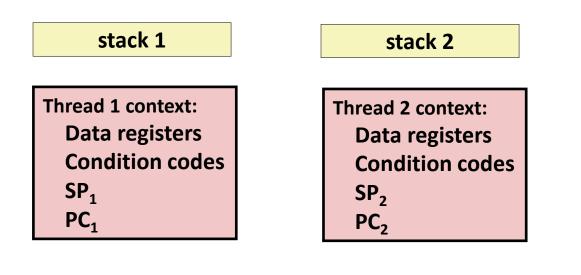
42

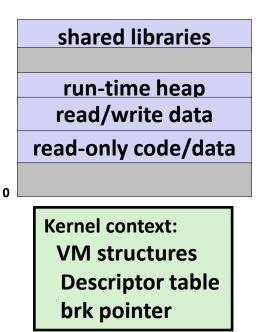
A Process With Multiple Threads

Multiple threads can be associated with a process

- Each thread has its own logical control flow
- Each thread shares the same code, data, and kernel context
- Each thread has its own stack for local variables
 - but not protected from other threads
- Each thread has its own thread id (TID)

Thread 1 (main thread) Thread 2 (peer thread)



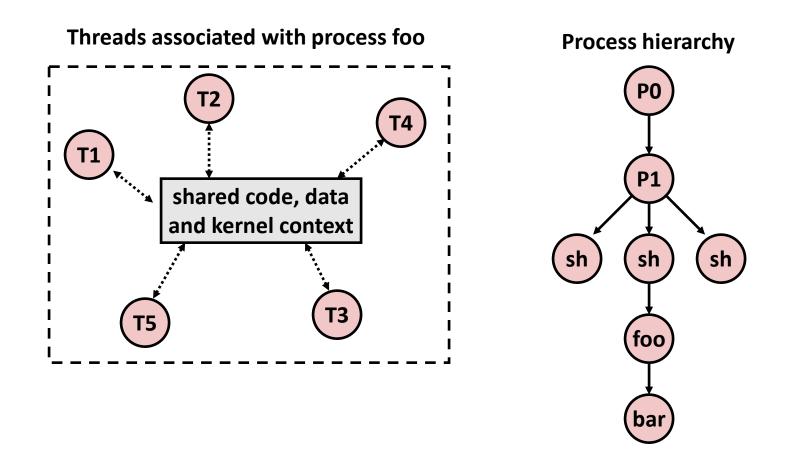


Shared code and data

Logical View of Threads

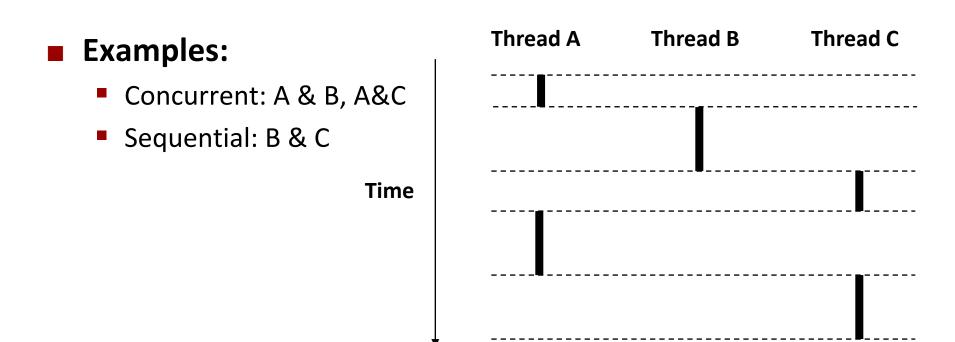
Threads associated with process form a pool of peers

Unlike processes which form a tree hierarchy



Concurrent Threads

- Two threads are concurrent if their flows overlap in time
- Otherwise, they are sequential



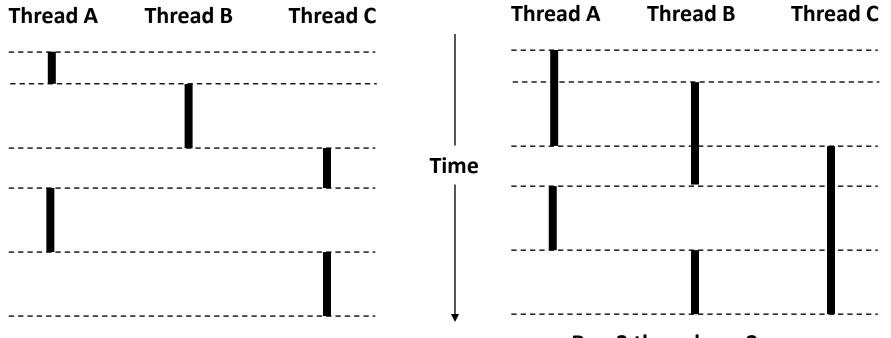
Concurrent Thread Execution

Single Core Processor

 Simulate parallelism by time slicing

Multi-Core Processor

 Can have true parallelism



Run 3 threads on 2 cores

Threads vs. Processes

How threads and processes are similar

- Each has its own logical control flow
- Each can run concurrently with others (possibly on different cores)
- Each is context switched

How threads and processes are different

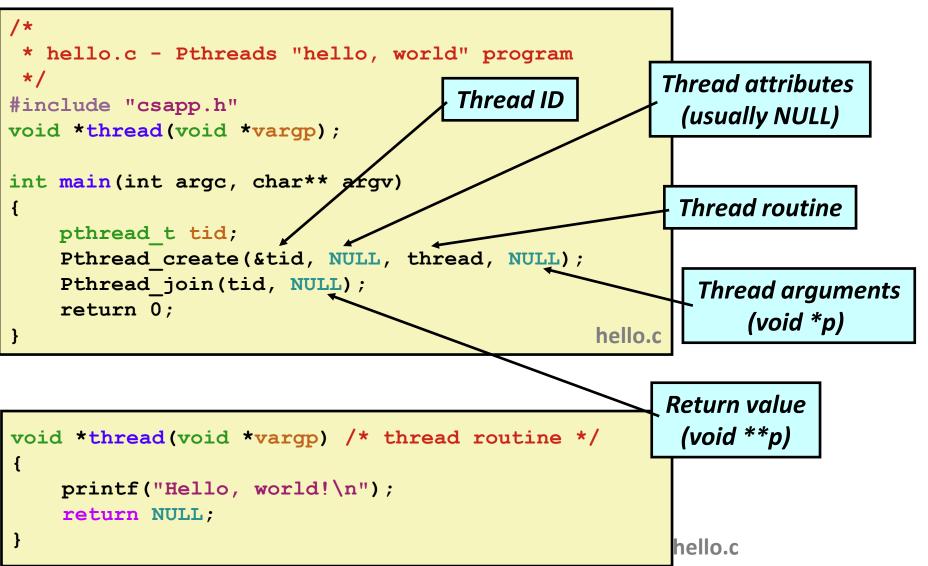
- Threads share all code and data (except local stacks)
 - Processes (typically) do not
- Threads are somewhat less expensive than processes
 - Process control (creating and reaping) twice as expensive as thread control
 - Linux numbers:
 - ~20K cycles to create and reap a process
 - ~10K cycles (or less) to create and reap a thread

Posix Threads (Pthreads) Interface

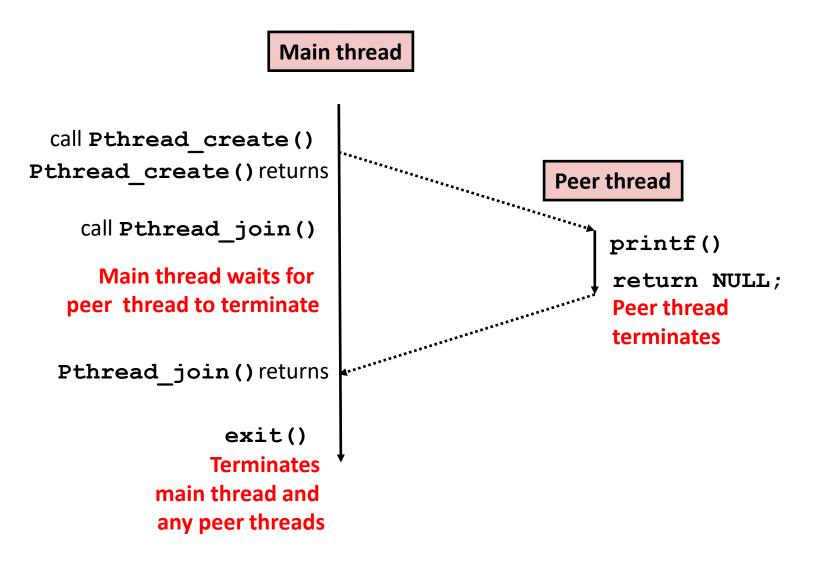
Pthreads: Standard interface for ~60 functions that manipulate threads from C programs

Thread API	Process API analogue	
Creating and reaping		
pthread_create	fork	
pthread_join	waitpid	
Determining your ID		
pthread_self	getpid	
Terminating execution		
pthread_exit	exit	
return from thread proc	return from main	
Synchronizing operations		
pthread_mutex_lock	[no exact analogue]	
pthread_mutex_unlock		

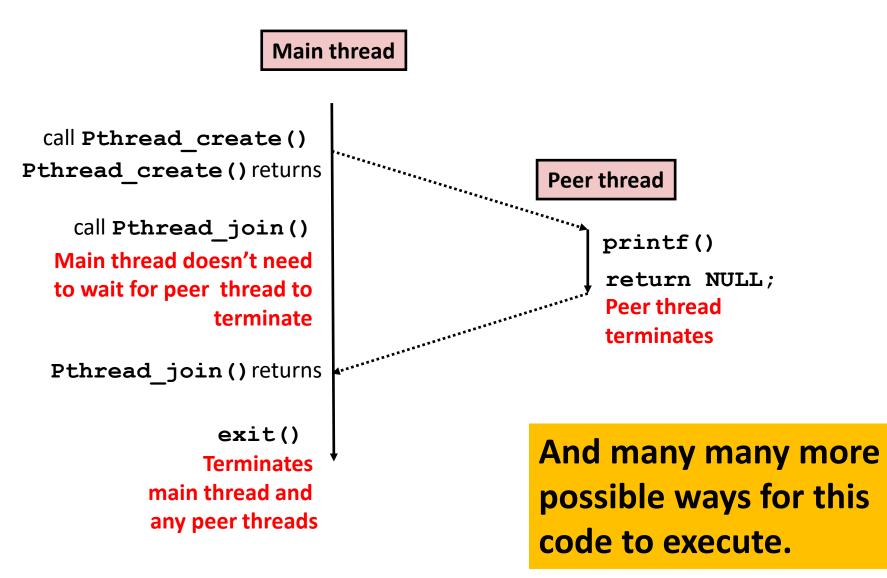
The Pthreads "hello, world" Program



Execution of Threaded "hello, world"



Or, ...



Thread-Based Concurrent Echo Server

```
int main(int argc, char **argv)
{
    int listenfd, *connfdp;
    socklen t clientlen;
    struct sockaddr storage clientaddr;
   pthread t tid;
    listenfd = Open listenfd(argv[1]);
   while (1) {
       clientlen=sizeof(struct sockaddr storage);
       connfdp = Malloc(sizeof(int));
       *connfdp = Accept(listenfd, (SA *) &clientaddr, &clientlen);
       Pthread create(&tid, NULL, thread, connfdp);
                                               echoservert.c
    return 0;
```

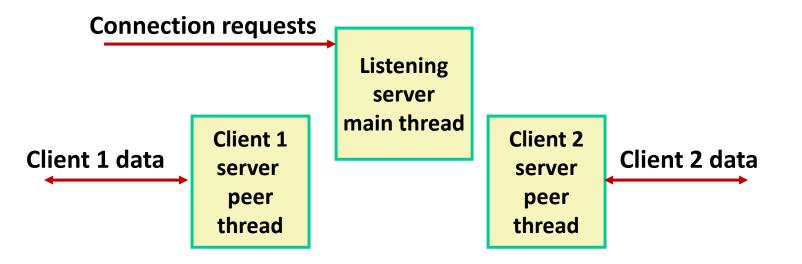
- Spawn new thread for each client
- Pass it copy of connection file descriptor
- Note use of Malloc() ! [but not Free()]

Thread-Based Concurrent Server (cont)

```
/* Thread routine */
void *thread(void *vargp)
{
    int connfd = *((int *)vargp);
    Pthread_detach(pthread_self());
    Free(vargp);
    echo(connfd);
    Close(connfd);
    return NULL;
} echoservert.c
```

- Run thread in "detached" mode.
 - Runs independently of other threads
 - Reaped automatically (by kernel) when it terminates
- Free storage allocated to hold connfd
- Close connfd (important!)

Thread-based Server Execution Model



- Each client handled by individual peer thread
- Threads share all process state except TID
- Each thread has a separate stack for local variables

Issues With Thread-Based Servers

Must run "detached" to avoid memory leak

- At any point in time, a thread is either *joinable* or *detached*
- Joinable thread can be reaped and killed by other threads
 - must be reaped (with pthread_join) to free memory resources
- Detached thread cannot be reaped or killed by other threads
 - resources are automatically reaped on termination
- Default state is joinable
 - use pthread_detach(pthread_self()) to make detached

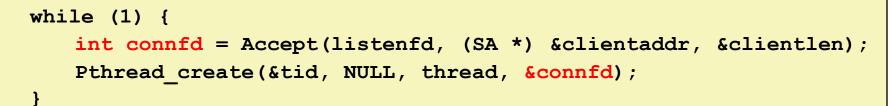
Must be careful to avoid unintended sharing

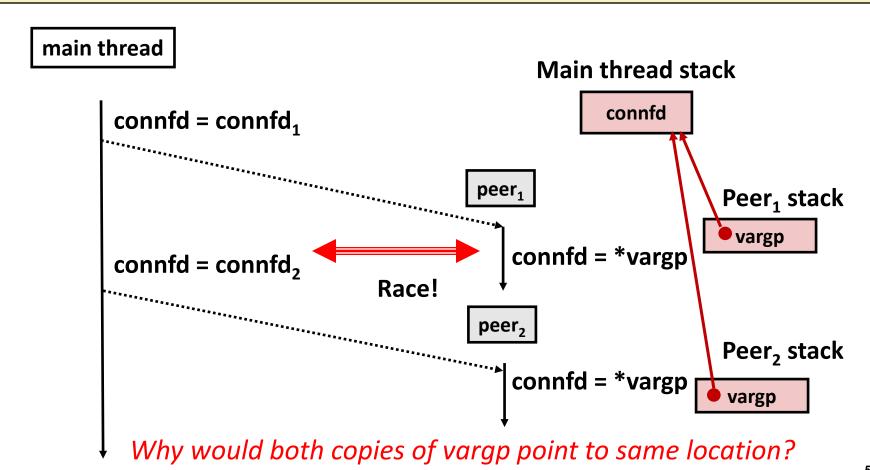
- For example, passing pointer to main thread's stack
 - Pthread_create(&tid, NULL, thread, (void *)&connfd);

All functions called by a thread must be thread-safe

(next lecture)

Potential Form of Unintended Sharing

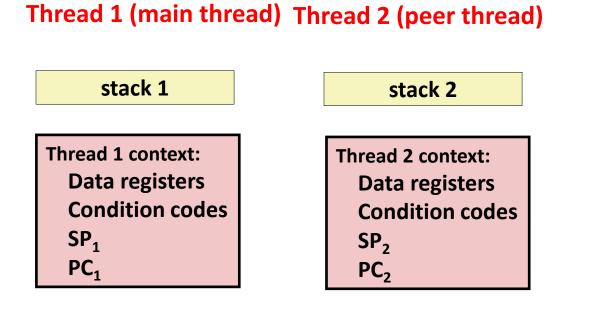




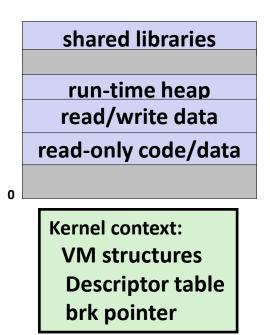
A Process With Multiple Threads

Multiple threads can be associated with a process

- Each thread has its own logical control flow
- Each thread shares the same code, data, and kernel context
- Each thread has its own stack for local variables
 - but not protected from other threads
- Each thread has its own thread id (TID)



Shared code and data



But ALL memory is shared

Thread 1 context: Data registers Condition codes SP₁ PC₁ Thread 2 context: Data registers Condition codes SP₂ PC₂

Thread 1 (main thread) Thread 2 (peer thread)

stack 1

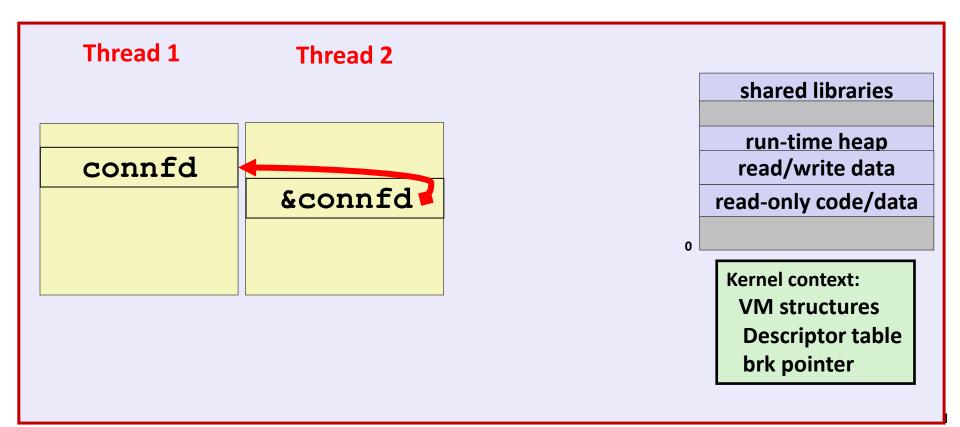
stack 2

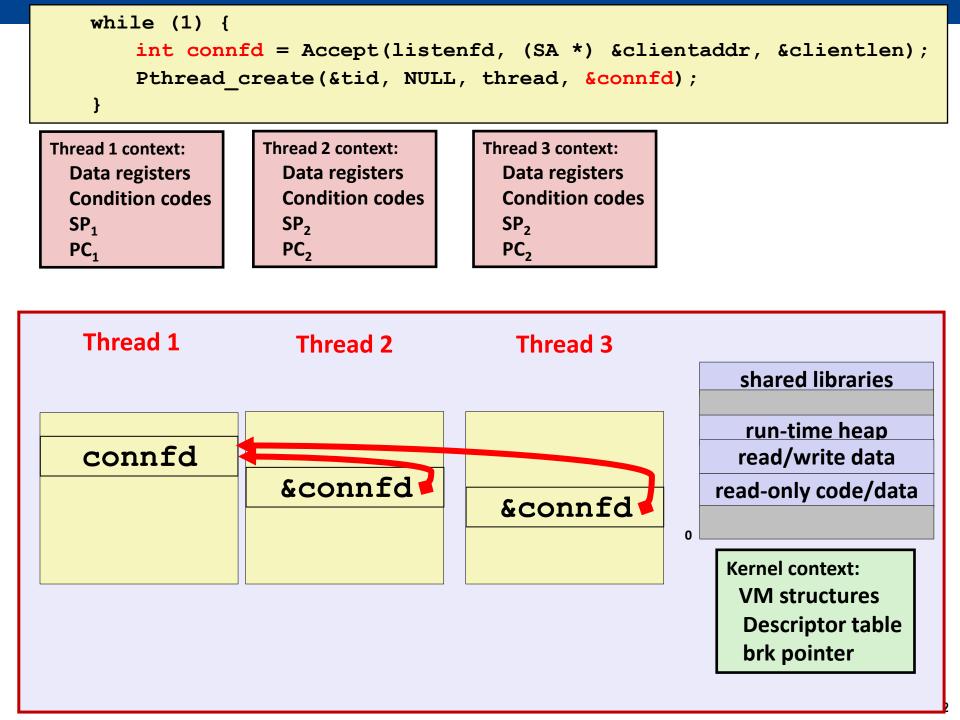
shared libraries run-time heap read/write data read-only code/data Kernel context: VM structures Descriptor table brk pointer

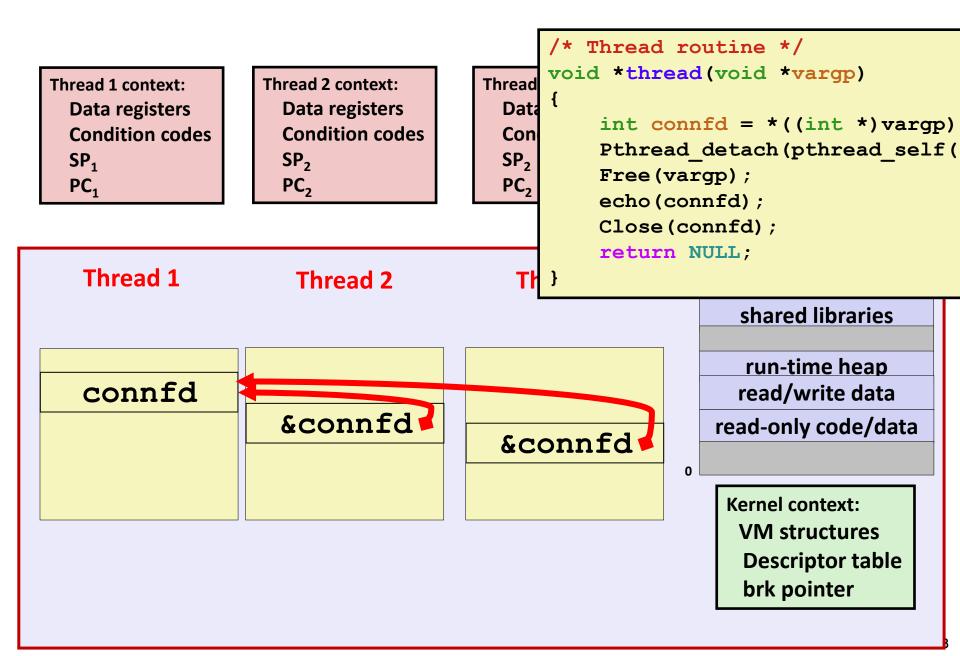
0

```
while (1) {
    int connfd = Accept(listenfd, (SA *) &clientaddr, &clientlen);
    Pthread_create(&tid, NULL, thread, &connfd);
}
```

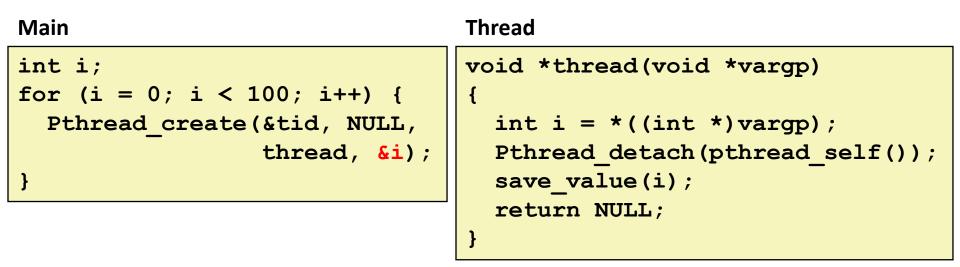
Thread 1 context:	Thread 2 context:
Data registers	Data registers
Condition codes	Condition codes
SP ₁	SP ₂
PC ₁	PC ₂







Could this race occur?

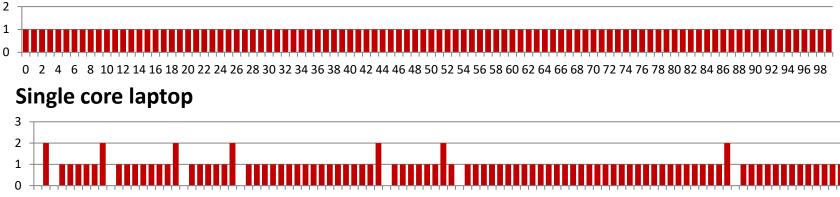


Race Test

- If no race, then each thread would get different value of i
- Set of saved values would consist of one copy each of 0 through 99

Experimental Results

No Race



0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94 96 98

Multicore server



Correct passing of thread arguments

```
/* Main routine */
    int *connfdp;
    connfdp = Malloc(sizeof(int));
    *connfdp = Accept( . . . );
    Pthread_create(&tid, NULL, thread, connfdp);
```

```
/* Thread routine */
void *thread(void *vargp)
{
    int connfd = *((int *)vargp);
    . . .
    Free(vargp);
    . . .
    return NULL;
}
```

Producer-Consumer Model

- Allocate in main
- Free in thread routine

Pros and Cons of Thread-Based Designs

+ Easy to share data structures between threads

- e.g., logging information, file cache
- + Threads are more efficient than processes
- Unintentional sharing can introduce subtle and hardto-reproduce errors!
 - The ease with which data can be shared is both the greatest strength and the greatest weakness of threads
 - Hard to know which data shared & which private
 - Hard to detect by testing
 - Probability of bad race outcome very low
 - But nonzero!
 - Future lectures

Summary: Approaches to Concurrency

Process-based

- Hard to share resources: Easy to avoid unintended sharing
- High overhead in adding/removing clients

Event-based

- Tedious and low level
- Total control over scheduling
- Very low overhead
- Cannot create as fine grained a level of concurrency
- Does not make use of multi-core

Thread-based

- Easy to share resources: Perhaps too easy
- Medium overhead
- Not much control over scheduling policies
- Difficult to debug
 - Event orderings not repeatable