CS318 Pintos Project Lab0 Overview

William Wang (xwill@bu.edu)

Outline

- Administrivia
- Lab overview
- Environment Setup
- Dev Tool
- Tips

Administrivia

Lab0 deadline: 09/19 23:59PM Friday, individually

score = code (70%) + design doc (30%)

Submit through GradeScope (only lab 0)

Outline

- Administrivia
- Lab overview
- Environment Setup
- Dev Tool
- Tips

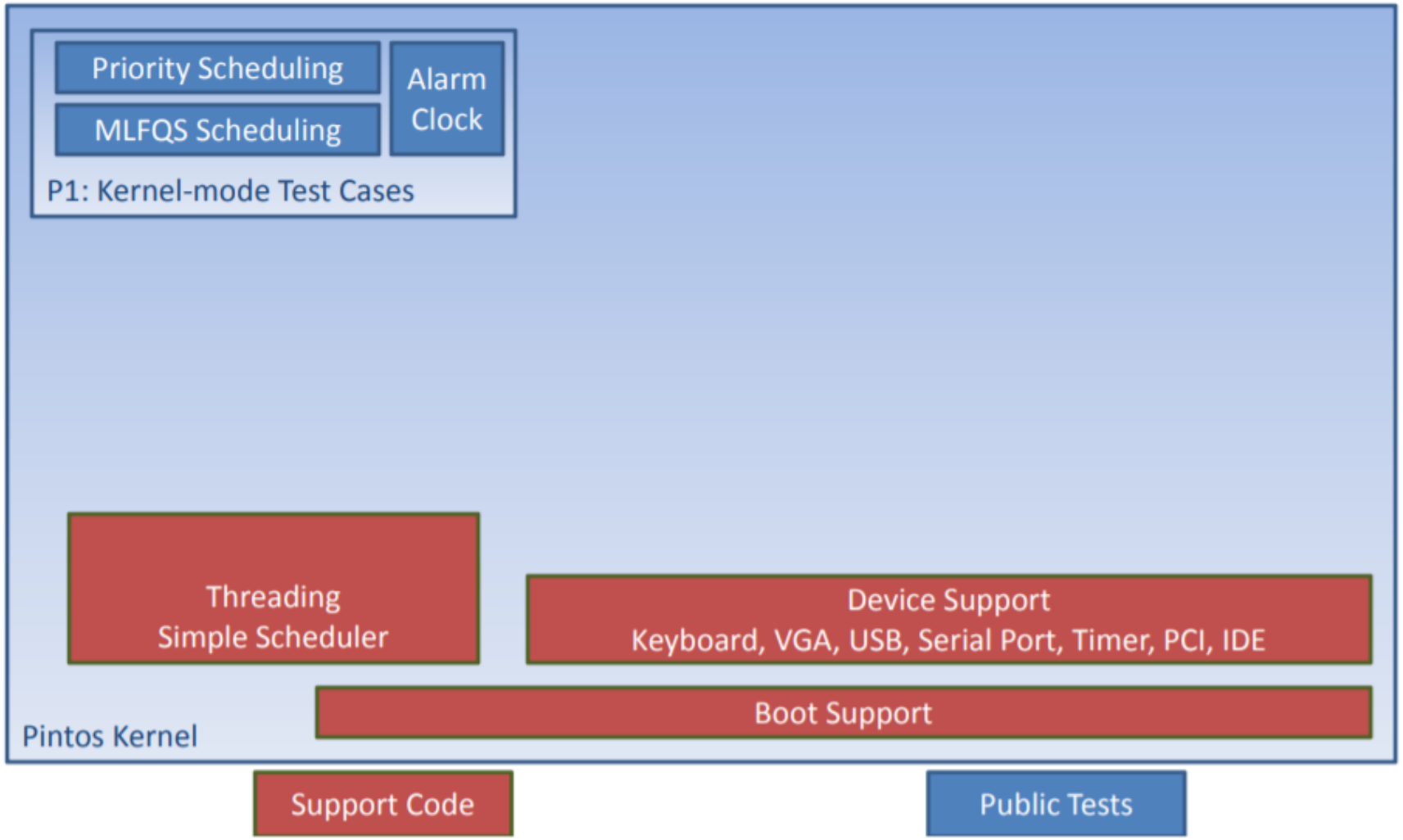
What Is Pintos?

- Pintos is a teaching operating system for 80x86
  - Developed in 2005 for Stanford's CS 140 OS class
  - Small enough so entire code can be read and understood by students

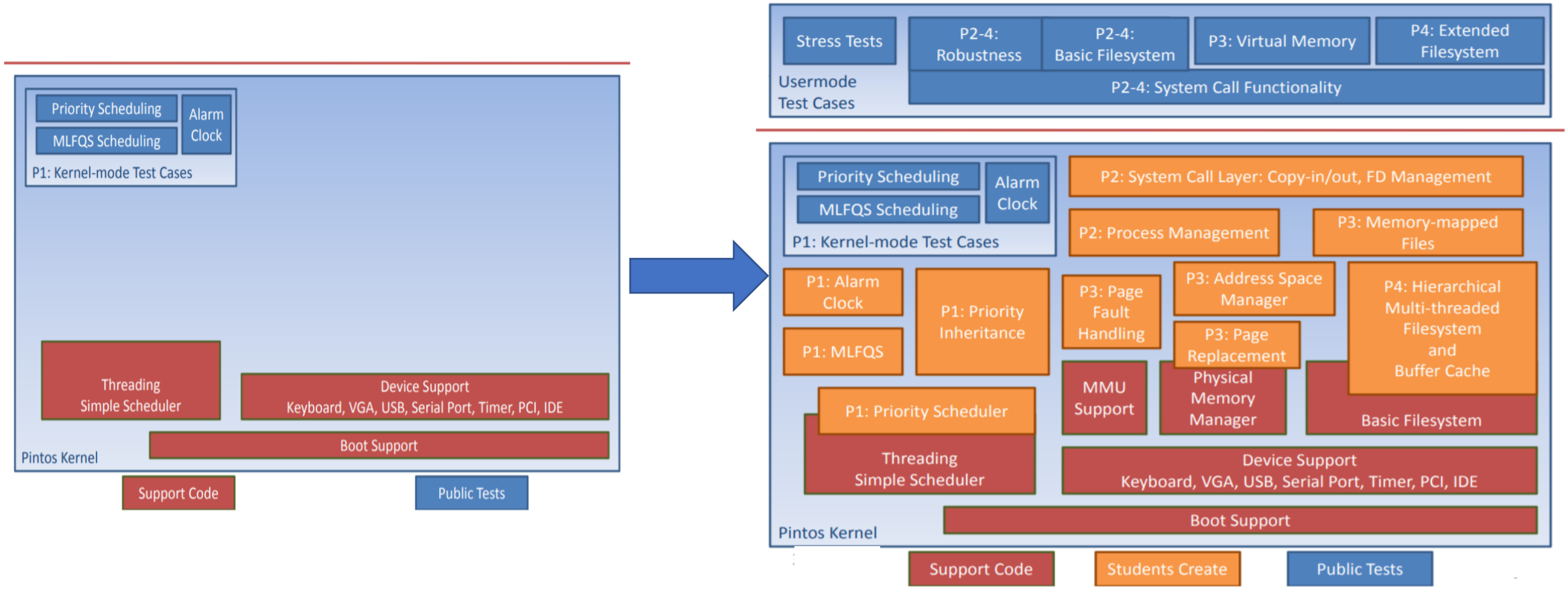
- Pintos supports kernel threads, virtual memory, user programs, and file system
  - Premature or incomplete

In this course project, you will improve all of these areas of Pintos to make it complete

Pintos Kernel(Pre Project)



Pintos Kernel(Post Project)



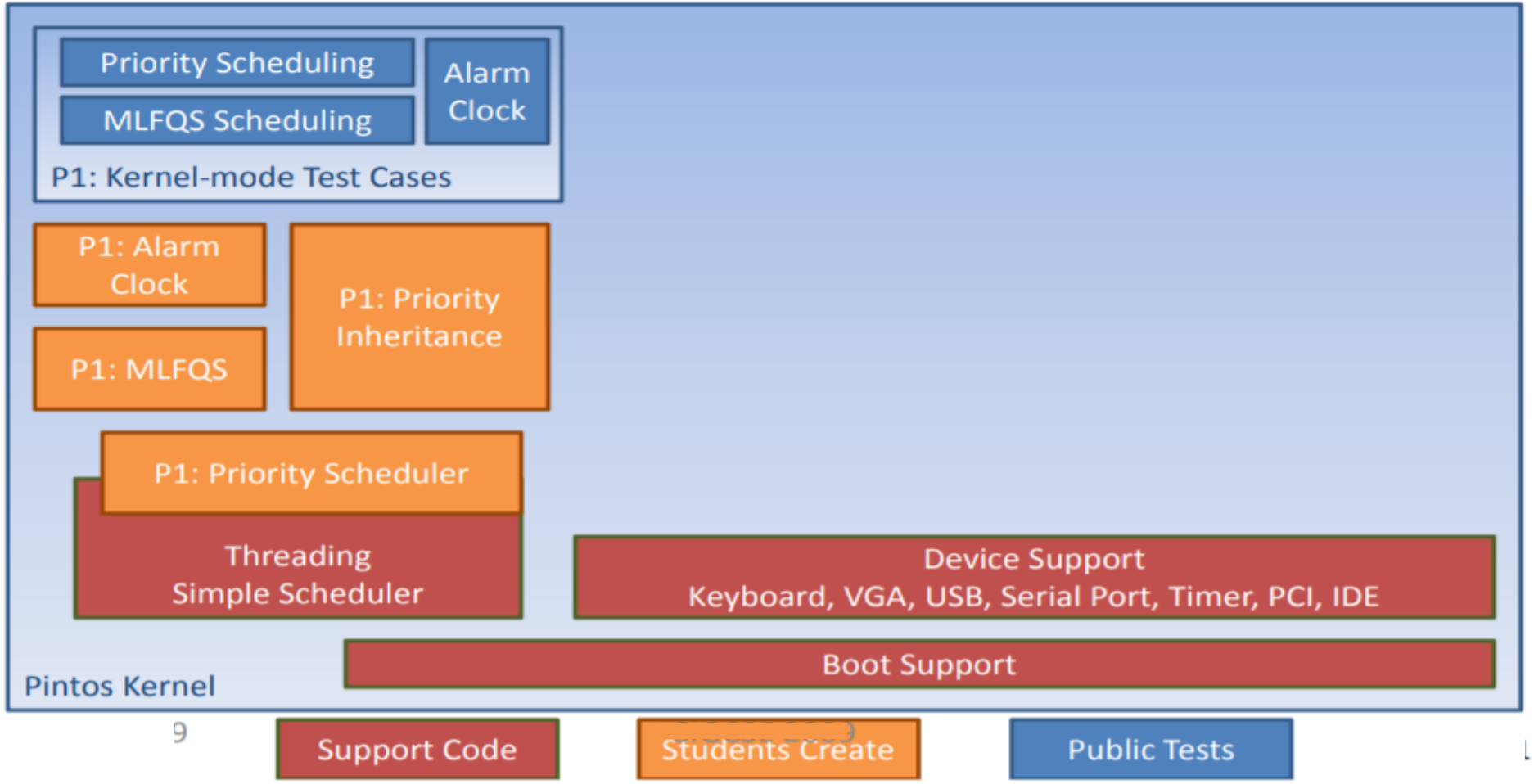
Lab 0 Overview

- Lab 0 is a warm-up exercise
  - Preparing you for the later Pintos projects
- In Lab 0, you will:
  - Install and boot Pintos
  - Go through the PC Bootstrap
  - Learn how to debug Pintos in QEMU and Bochs
  - Add a tiny kernel monitor to Pintos

https://yigonghu.github.io/EC440/fall25/projects/lab0

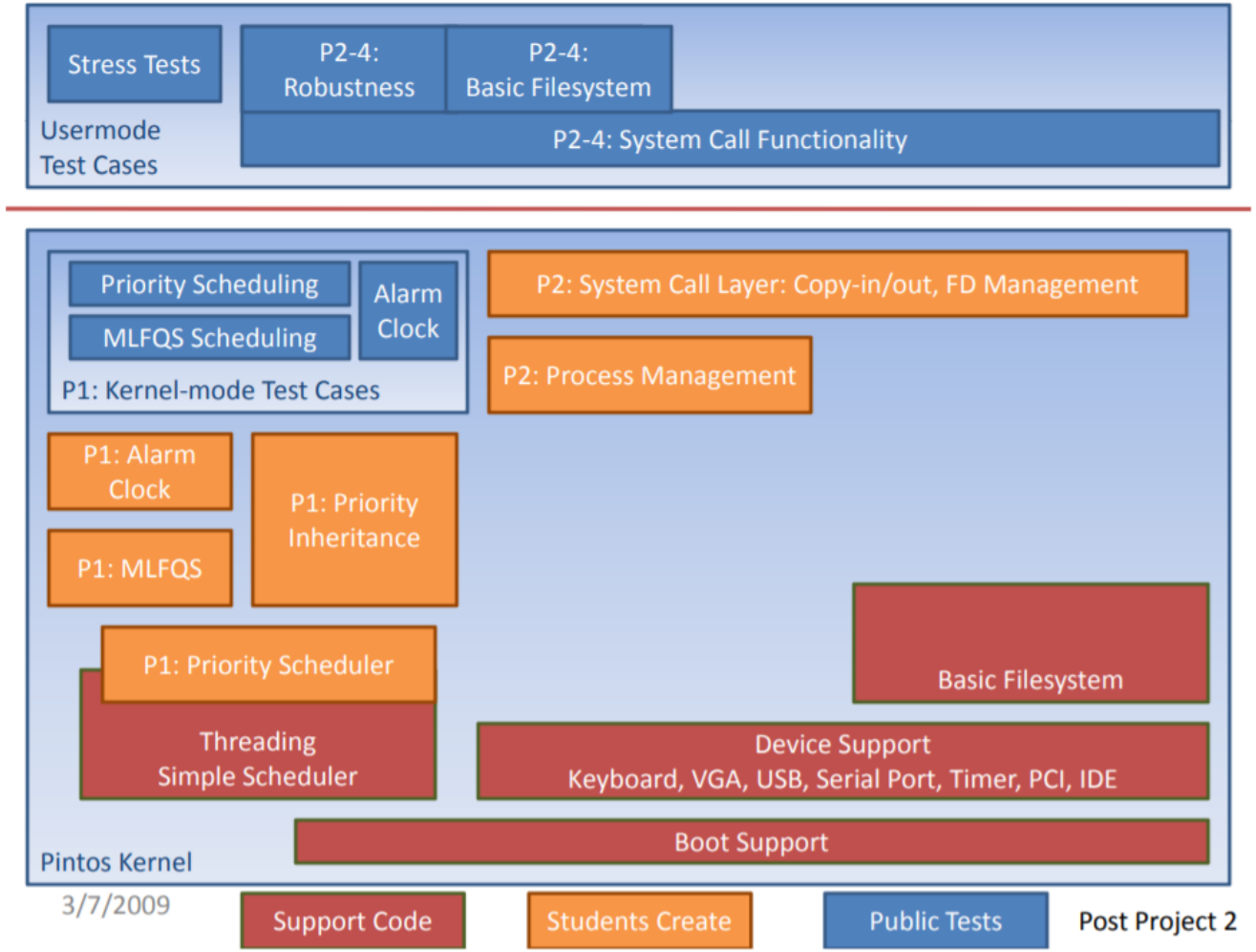
Lab 1 Overview

- Extending the functionality of Pintos thread system



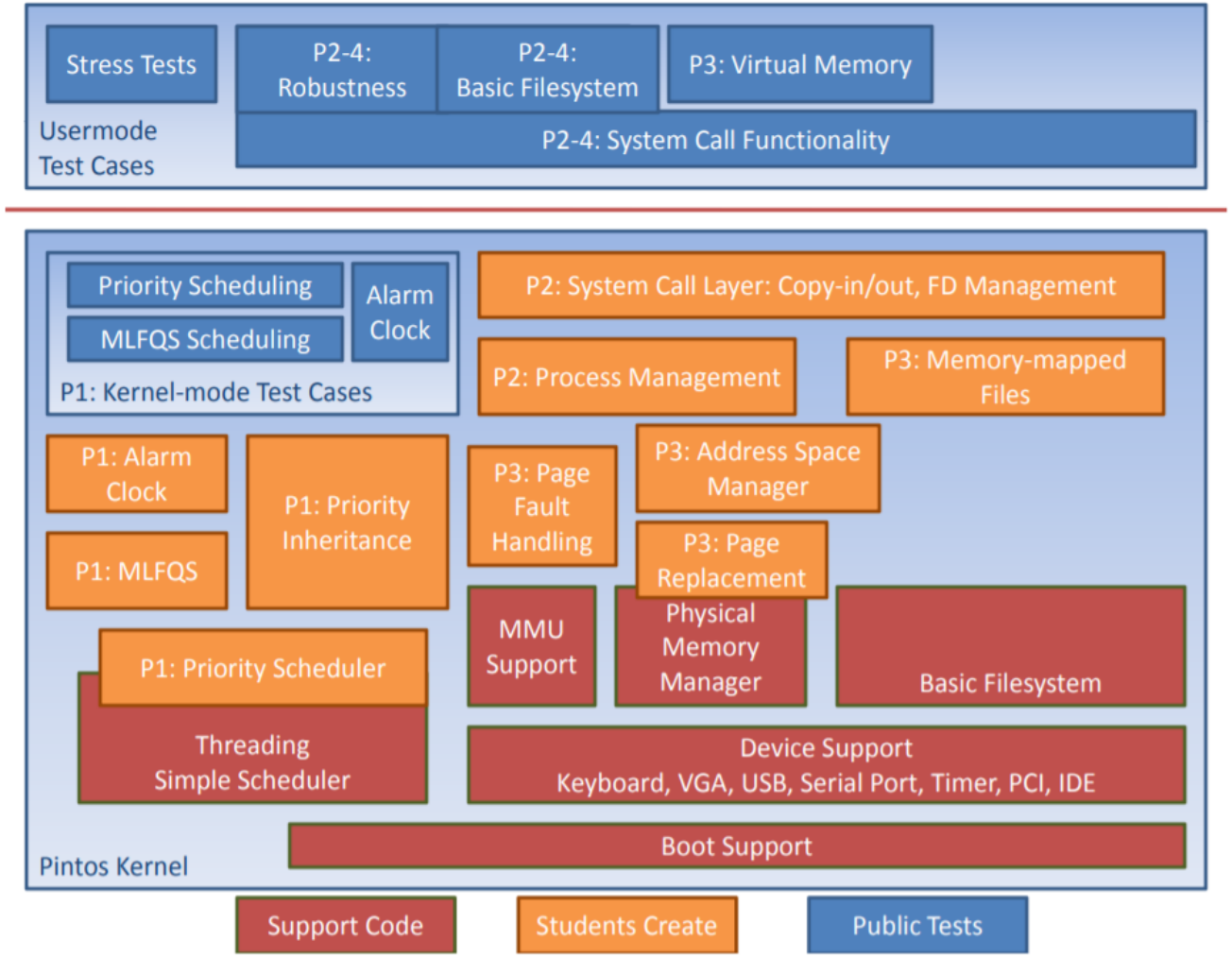
Lab 2 Overview

- Enable user programs to interact with the OS via system calls



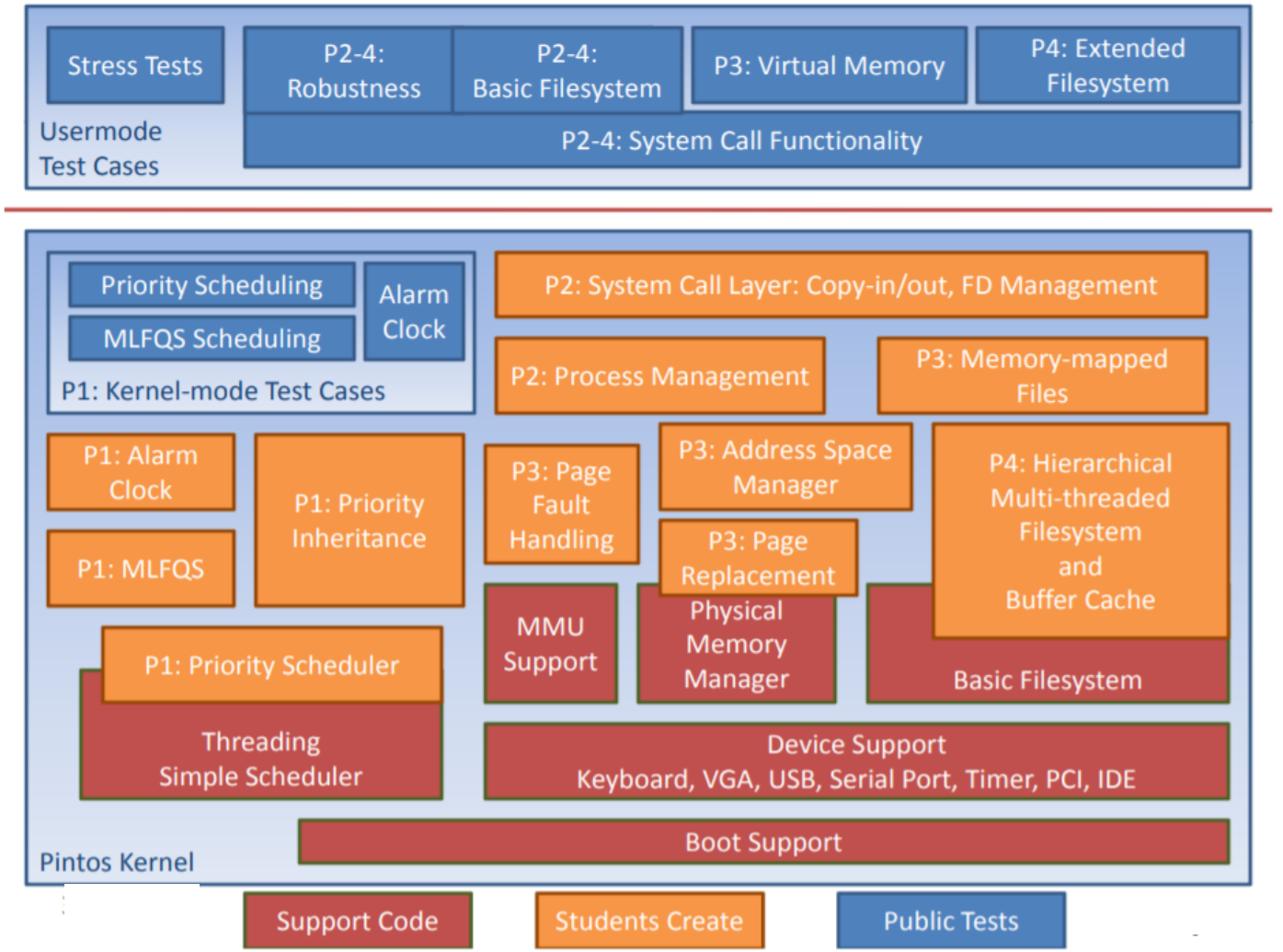
Lab 3 Overview

- Implement the virtual memory management

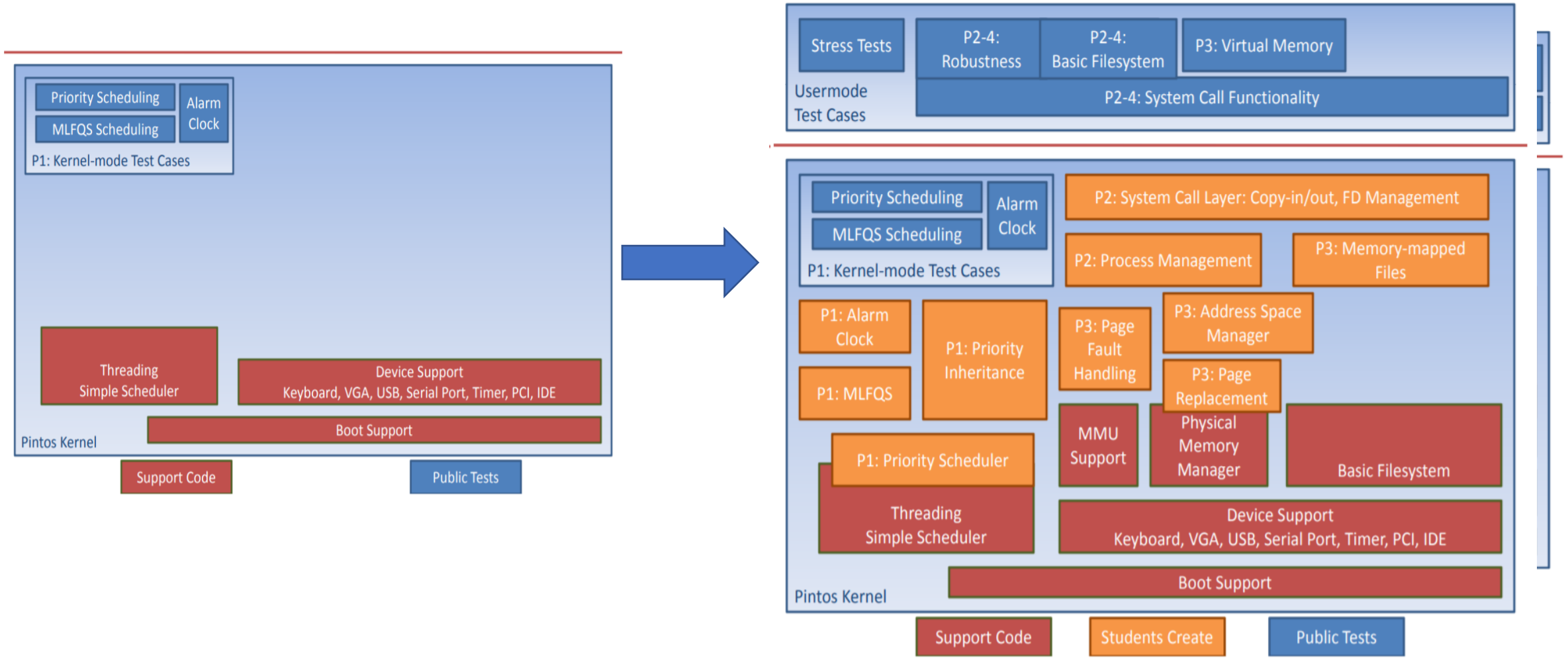


Lab 4 Overview

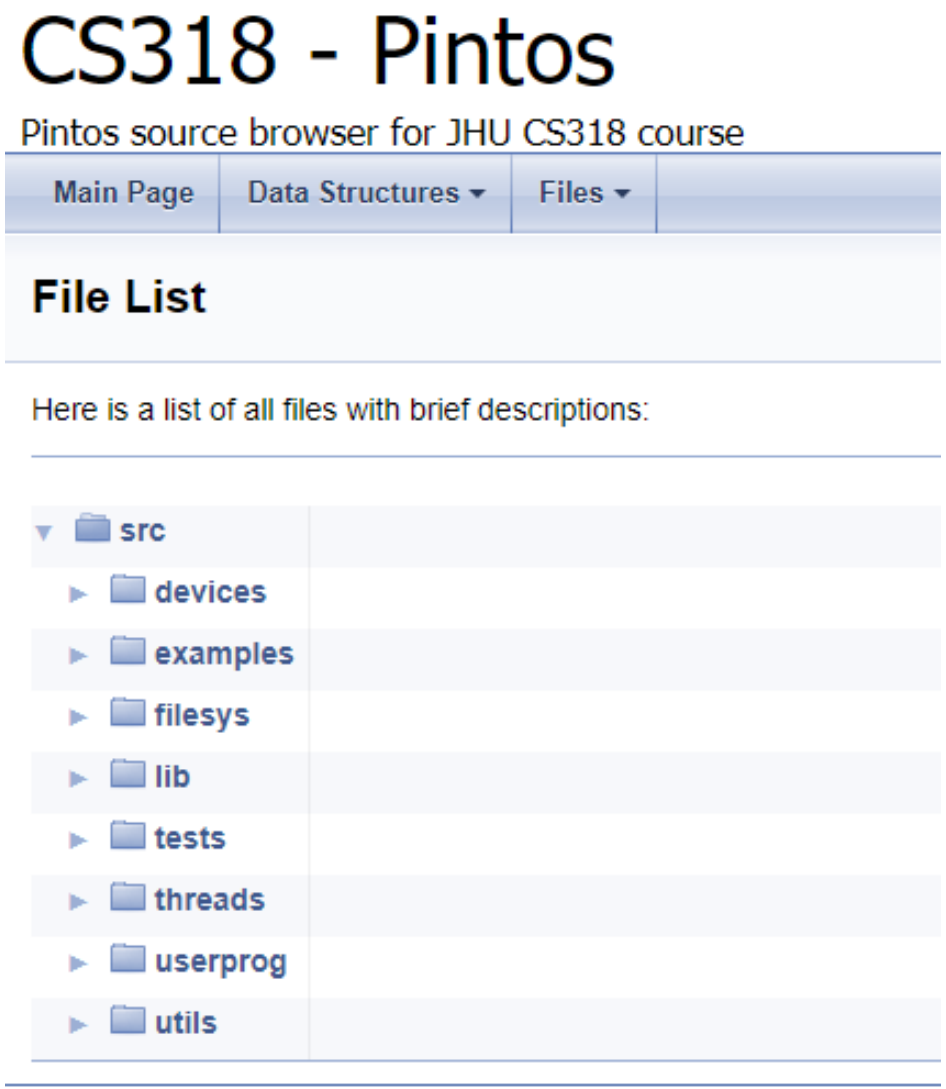
- Extending basic filesystem to hierarchical filesystem



Pintos Kernel(Post Project)



Pintos Source Code Overview



https://jhu-cs318.github.io/pintos-doxygen/html/files.html

Pintos Source Tree

- threads/
  - Source code for the base kernel, which you will modify starting in project 1
- userprog/
  - Source code for the user program loader, which you will modify starting with project 2
- vm/
  - An almost empty directory. You will implement virtual memory here in project 3.
- filesys/
  - Source code for a basic file system. You will use this file system starting with project 2, but you will not modify it until project 4

Pintos Source Tree

- devices/
  - Source code for I/O device interfacing: keyboard, timer, disk, etc. You will modify the timer implementation in project 1.
- lib/
  - An implementation of a subset of the standard C library.
- tests/
  - All the test cases for each project
- examples/
  - Example user programs for use starting with project 2

Pintos Source Tree

- misc/
- utils/
  - These files may come in handy if you decide to try working with Pintos on your own machine. Otherwise, you can ignore them

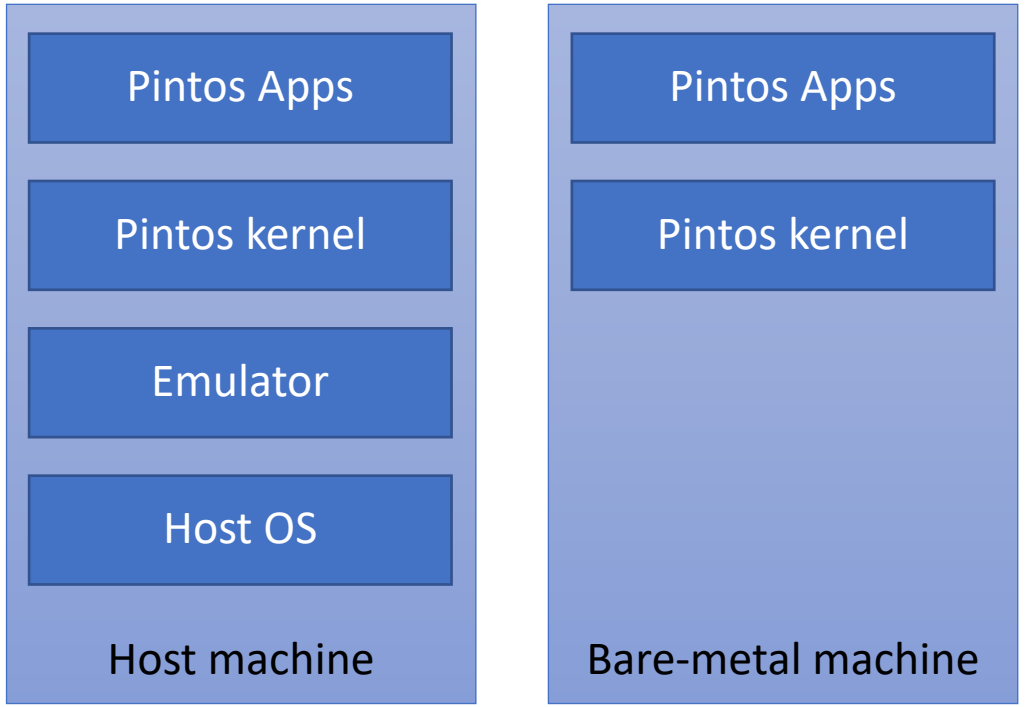
Outline

- Administrivia
- Lab 0 overview
- Environment Setup
- Dev Tool
- Tips

Pintos Running Environment

- Pintos can run and debug in
  - Emulated environment
  - Real hardware

In this course, we will use emulated environment.



Environment

- Required Tools for PintOS
  - 80x86 cross-compiler toolchain for 32-bit architecture
    - ➢ C compiler, assembler, linker, and debugger.
  - x86 emulator
    - ➢ QEMU or Bochs

- Working Environment
  - SCC lab machine
  - Pre-made docker/utm image
  - Your own machine

Use premade docker/virtual machine image

- Premade containers already have toolchains installed
- Environments are setup in advance, use out-of-the-box
- Useful Resource
  - Docker image
    docker run -it —name pintos buec440/pintos bash
  - UTM image
    To be added

Installing Pintos on You Own Machine

- You may want to work on your own machines to be more productive
- Need to build the toolchain
- Useful Links
  - Build script (tested on Mac, Ubuntu and Fedora)
    pintos/src/misc/toolchain-build.sh
  - Installation guide
    https://yigonghu.github.io/EC440/fall25/projects/setup

Environment setting on SCC Lab Machine

- CS lab machines already have required tools installed
- To include tools, you need source the ec440 tools
  - Change working directory into /projectnb/ec440/projects and:

    source envsetup.sh
- Build pintos
  - git clone https://github.com/yigonghu/ec440-pintos.git
  - cd pintos/src/threads
  - make

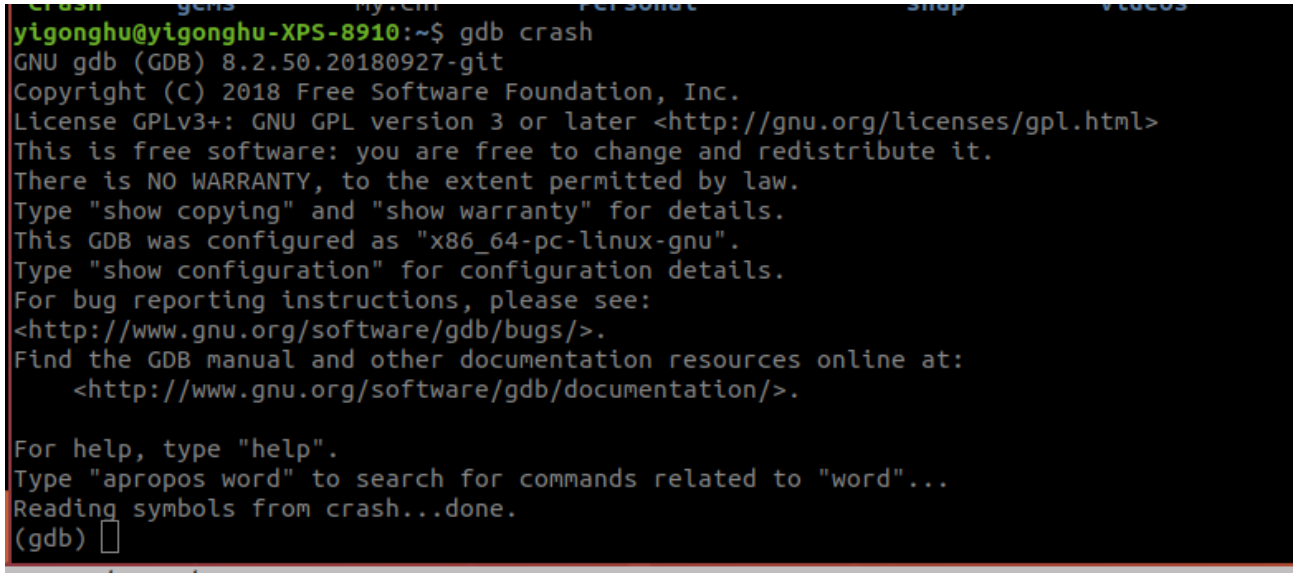
Outline

- Administrivia
- Lab 0 overview
- Environment Setup
- Dev Tool
- Tips

GDB

- “GNU Debugger”
- A debugger for several languages, including C and C++
- The Pintos uses GDB as the default debugging tool.
- Online manual

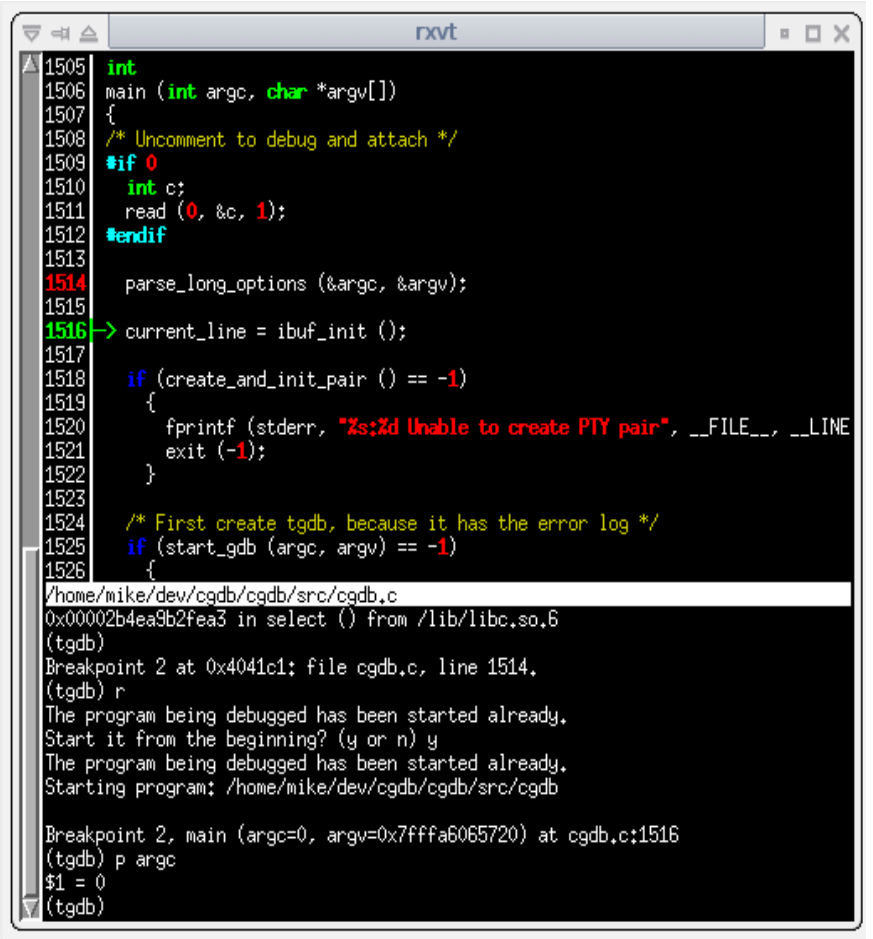
  https://sourceware.org/gdb/current/onlinedocs/gdb/



CGDB

- CGDB
  - A lightweight curses interface to GDB
  - Standard gdb console
  - A split screen view that displays the source code
- Strongly recommend using CGDB to debug Pintos
- Reference

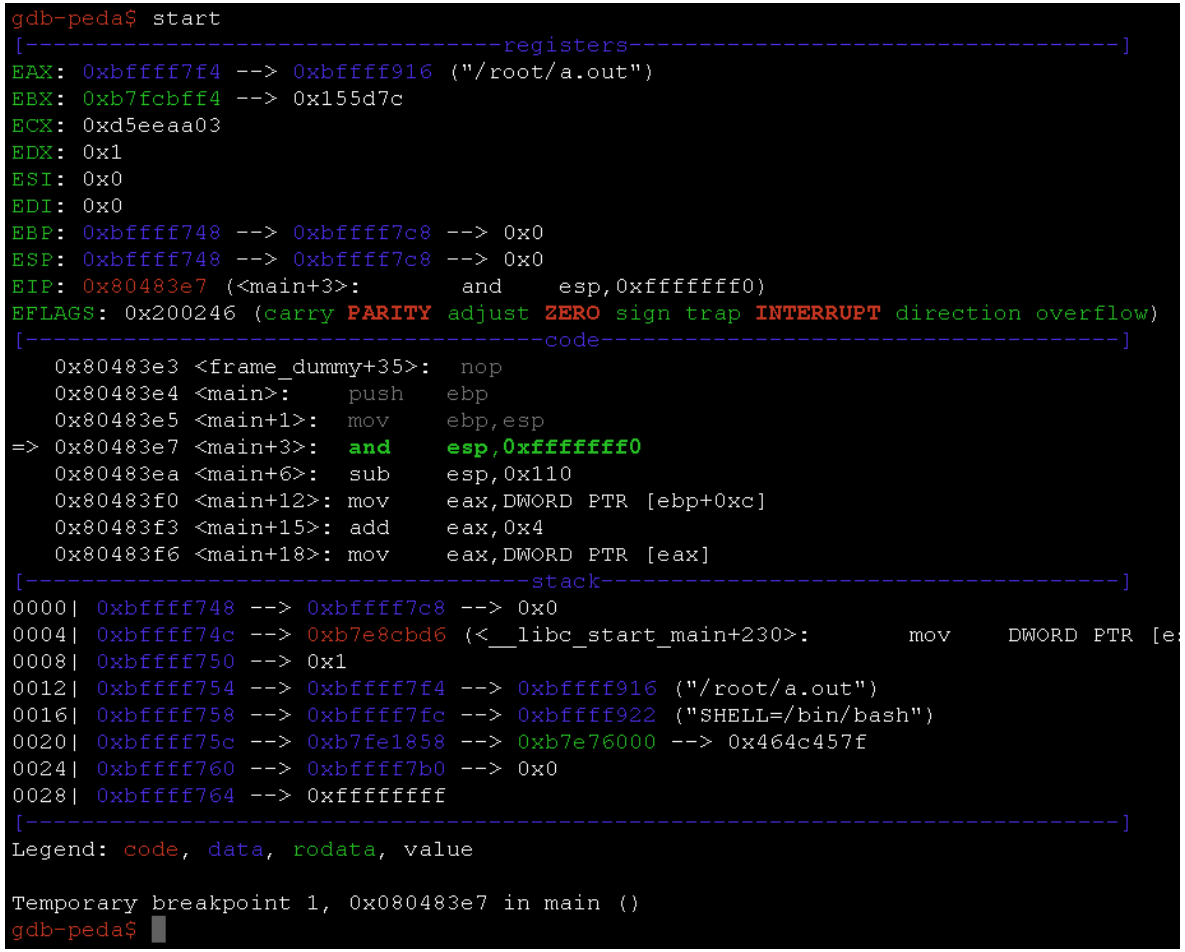
  https://cgdb.github.io/



PEDA

- PEDA
  - An extension for GDB written in Python
  - Standard gdb console
  - A split screen view that displays registers, assembly and stack

- Reference
  https://github.com/longld/peda



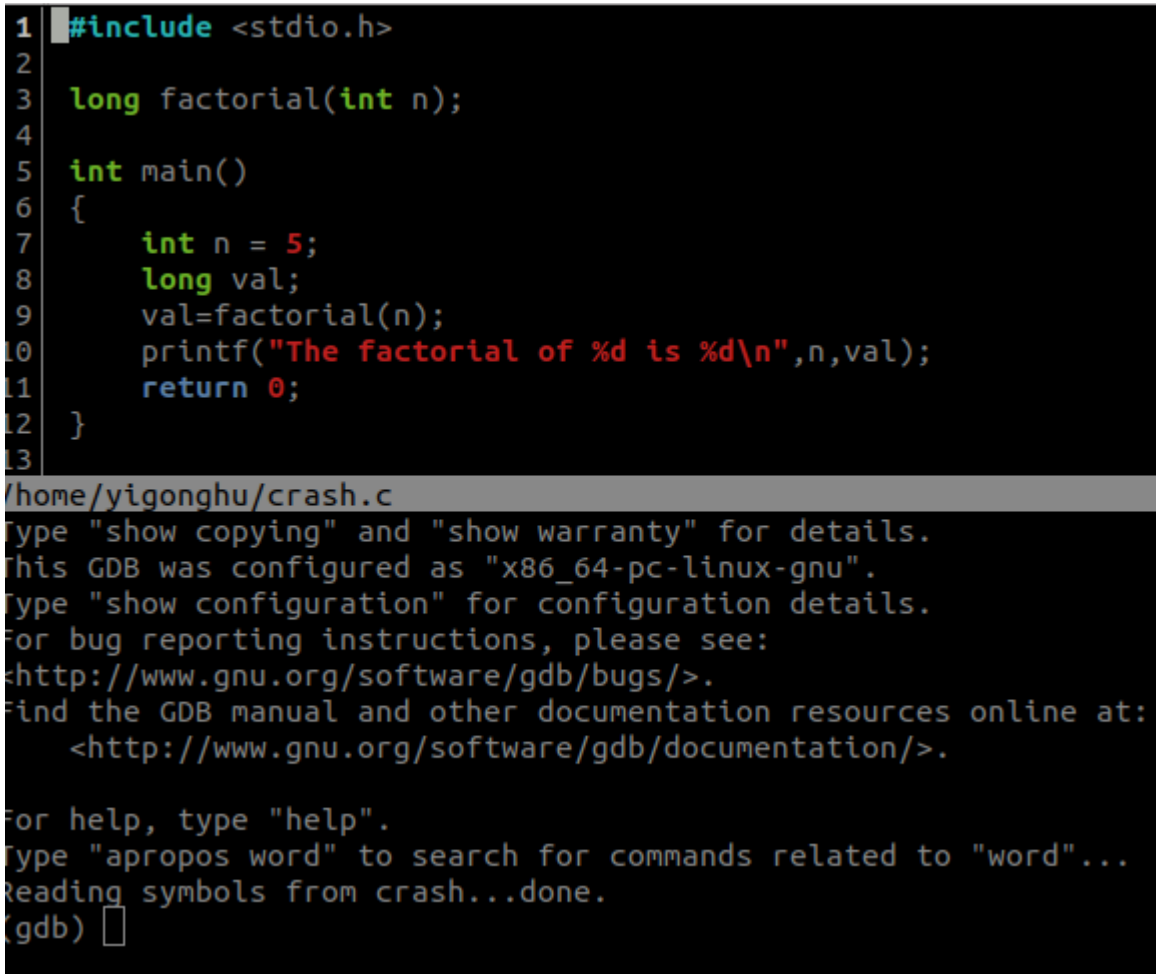
How to Use CGDB to Debug A Toy Program?

```c
#include <stdio.h>

long factorial(int n);

int main()
{
    int n = 5;
    long val;
    val=factorial(n);
    printf("The factorial of %d is %d\n",n,val);
    return 0;
}

long factorial(int n)
{
    long result = 1;
    while(n--)
    {
        result*=n;
    }
    return result;
}
```

The given code computes the factorial of a number erroneously. The program always outputs 0, regardless of the input

Starting Up CGDB

- To run the cgdb, just first try “cgdb <filename>” (build with `gcc -g`)



Running Program

- To run the buggy program, use:

  (gdb) run

- This runs the program
  - ➢ If it has no serious problems, the program should run fine here too.
  - ➢ If the program did have issues, then you should get some useful information like the line number where it crashed, and parameters to the function that caused the error

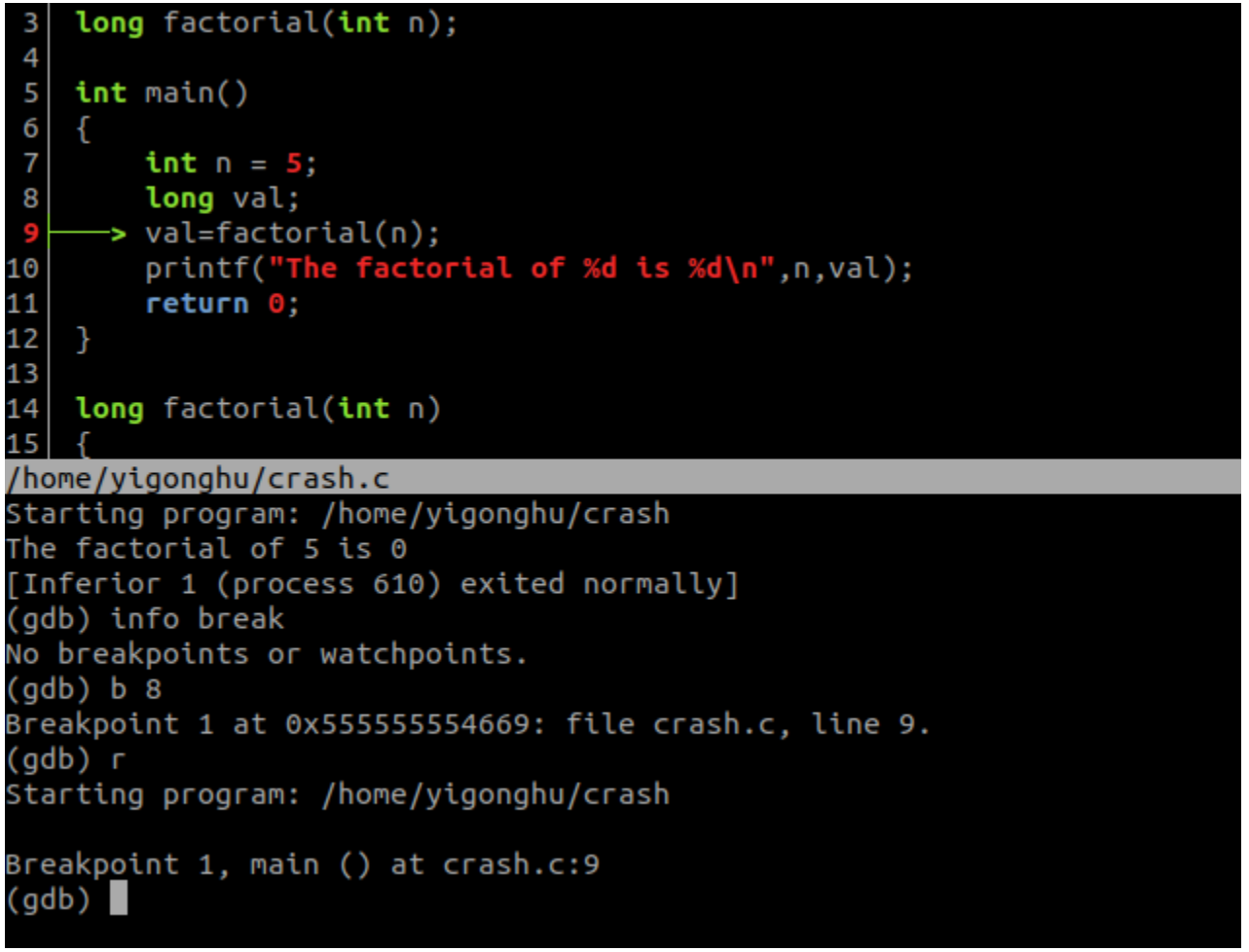
Setting Breakpoints

- What if the program output doesn’t match your expectation?
  - Step through your code a bit at a time, until you arrive upon the error
- Breakpoints can be used to stop the program run in the middle, at a designated point.

```
1  #include <stdio.h>
2
3  long factorial(int n);
4
5  int main()
6  {
7      int n = 5;
8      long val;
9      val=factorial(n);
10     printf("The factorial of %d is %d\n",n,val);
11     return 0;
12 }
13
/home/yigonghu/crash.c
For help, type "help".
Type "apropos word" to search for commands related to "word"...
Reading symbols from crash...done.
(gdb) r
Starting program: /home/yigonghu/crash
The factorial of 5 is 0
[Inferior 1 (process 610) exited normally]
(gdb) info break
No breakpoints or watchpoints.
(gdb) b 8
Breakpoint 1 at 0x555555554669: file crash.c, line 9.
(gdb)
```

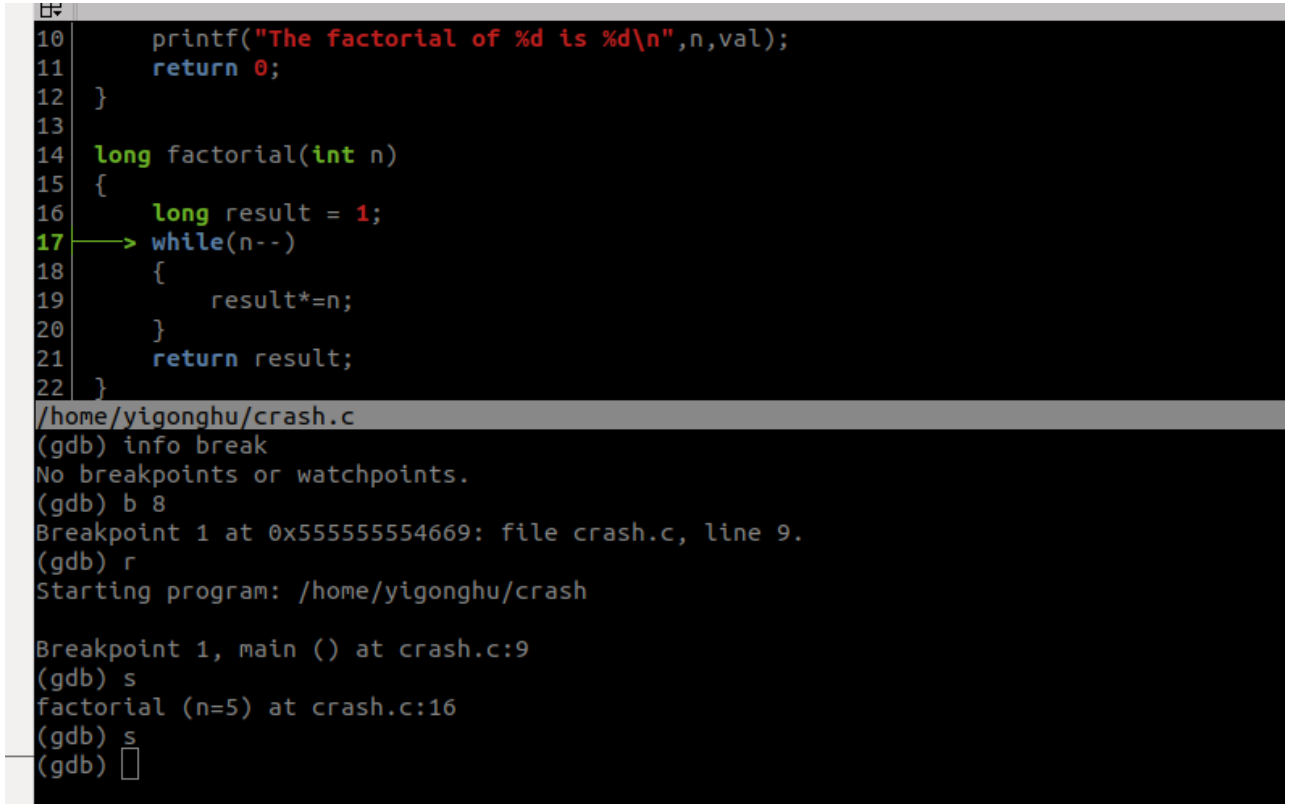
Run Program with breakpoints

- Once you've set a breakpoint, you can try using the run command again, This time, it should stop where you tell it to.



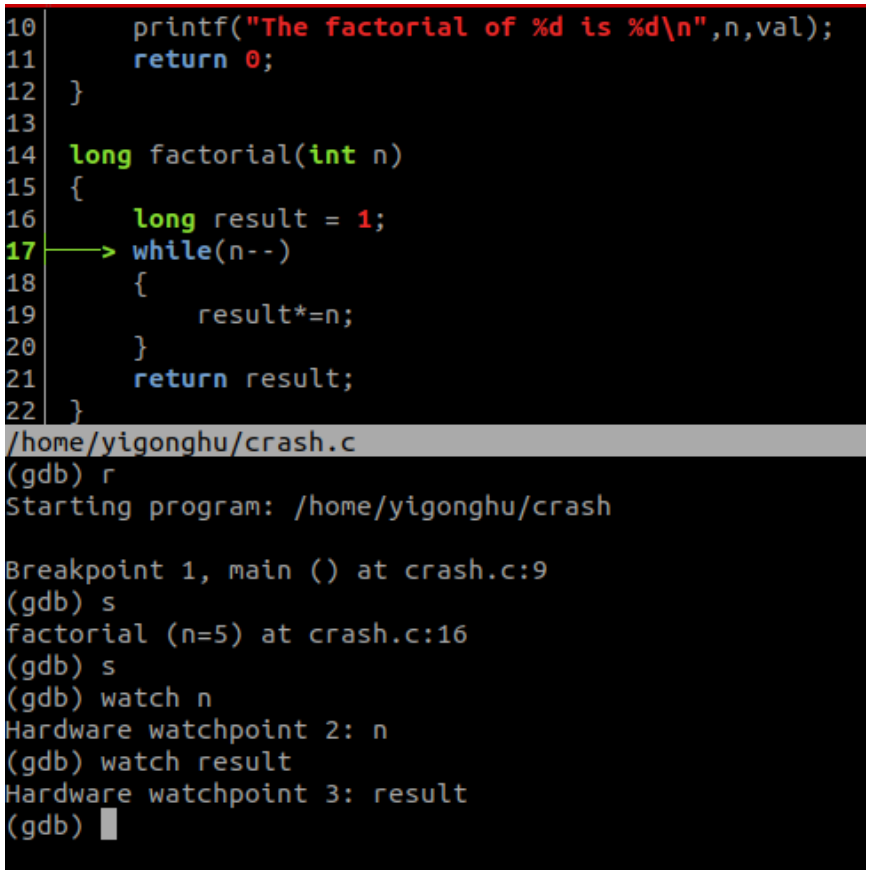
Step Into Code

- You can single-step (execute just the next line of code) by typing “step.” This gives you really fine-grained control over how the program proceeds. You can do this a lot.



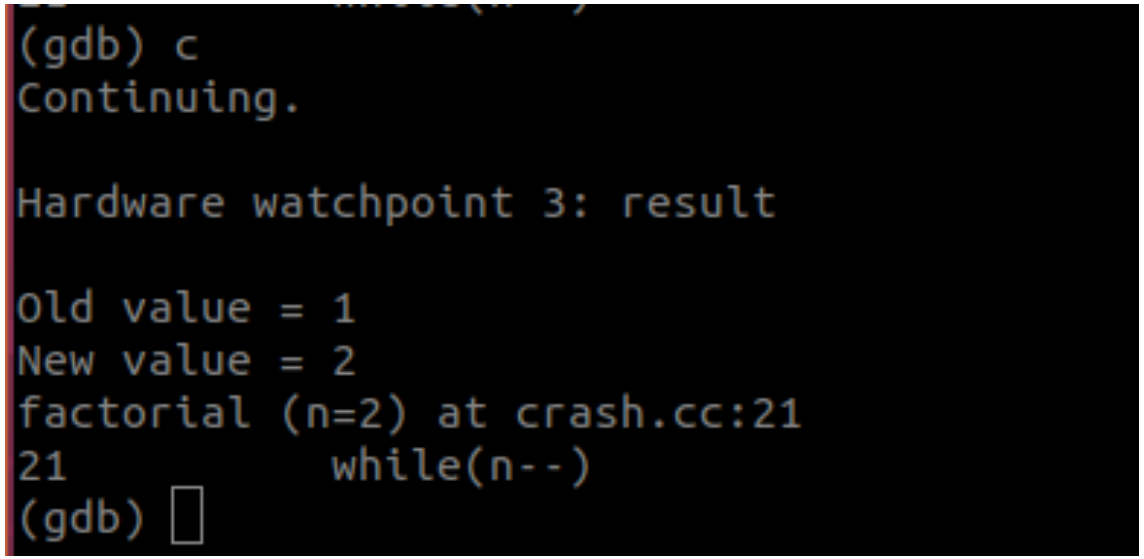
How to Check The Variable Change?

- Setting Watchpoints
  - Watchpoints interrupt the program whenever a watched variable's value is modified



Checking the Value of Variable

- After running the program, We've found the first bug! result is supposed to be evaluated by multiplying 3 * 2 * 1 but here the multiplication starts from 2. To correct it, we have to change the loop a little bit, but before that, lets see if the rest of the calculation is correct



GDB on Pintos (1)

- Pintos uses gdb's remote debugging feature

- Debugging Pintos is a two-step process
  - Starting pintos with gdb option: pintos --gdb -- run mytest
  - A second terminal to invoke GDB on kernel image: pintos-gdb kernel.o
  - Within the GDB shell, issue command: target remote localhost:1234
    - short-hand for the command: debugpintos

- If using lab machine, you need to use a custom port to avoid port conflict error: --gdb-port = port_number
  - Use the same port number in GDB: target remote localhost:port_number

GDB on Pintos (2)

- CGDB
  - The CS lab machine has cgdb installed
  - pintos-gdb will automatically prefer cgdb if it's available

Debugging On Pintos

- We introduce a bug into Pintos by commenting the initialization of ready_list, which will cause an assertion error in Pintos

```
85    It is not safe to call thread_current() until this function
86    finishes. */
87 void
88 thread_init (void)
89 {  ASSERT (intr_get_level () == INTR_OFF);
90
91   lock_init (&tid_lock);
92  // list_init (&ready_list);
93   list_init (&all_list);
94
95   /* Set up a thread structure for the running thread. */
96   initial_thread = running_thread ();
97   init_thread (initial_thread, "main", PRI_DEFAULT);
98   initial_thread->status = THREAD_RUNNING;
99   initial_thread->tid = allocate_tid ();
100 }
```

Debugging On Pintos

Machine View

```
SeaBIOS (version ?-20180724_192412-buildhw-07.phx2.fedoraproject.org-1.fc29)
Booting from Hard Disk...
Pintos hda1
Loading.............
Kernel command line: run alarm-zero
Pintos booting with 3,968 kB RAM...
367 pages available in kernel pool.
367 pages available in user pool.
Kernel PANIC at ../../lib/kernel/list.c:171 in list_insert(): assertion `is_inte
rior (before) || is_tail (before)' failed.
Call stack: 0xc0029717 0xc0029b7e 0xc0029d22 0xc0020c64 0xc0020b6a 0xc00209a3 0x
c0020372.
The `backtrace' program can make call stacks useful.
Read "Backtraces" in the "Debugging Tools" chapter
of the Pintos documentation for more information.
```

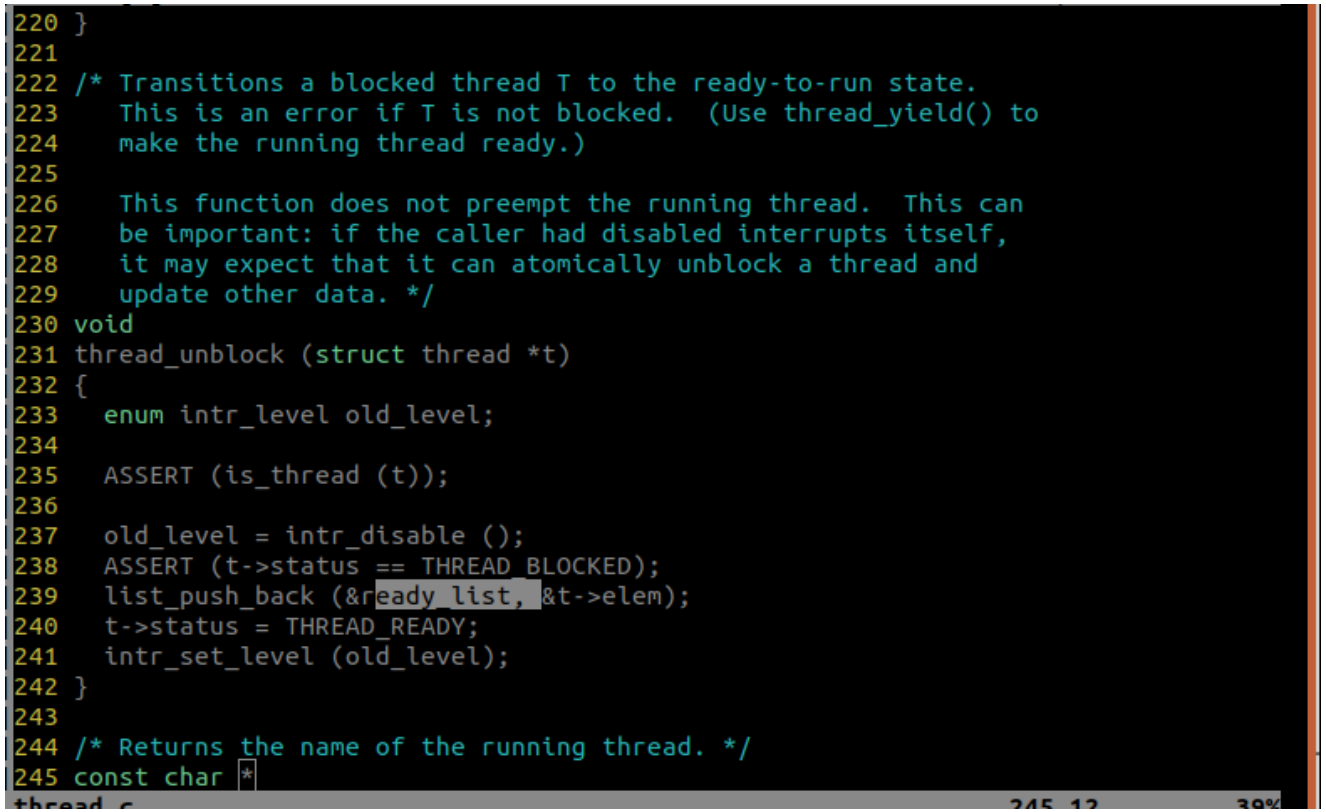
Debugging On Pintos

- To trace the call stack that cause the failure, we set a break point in list.c:170 and run the Pintos

```
159  list_tail (struct list *list)
160  {
161    ASSERT (list != NULL);
162    return &list->tail;
163  }
164
165  /* Inserts ELEM just before BEFORE, which may be either an
166     interior element or a tail.  The latter case is equivalent to
167     list_push_back(). */
168  void
169  list_insert (struct list_elem *before, struct list_elem *elem)
170 ->{
171    ASSERT (is_interior (before) || is_tail (before));
172    ASSERT (elem != NULL);
173
174    elem->prev = before->prev;
175    elem->next = before;
176    before->prev->next = elem;
177    before->prev = elem;
178  }
179
180  /* Removes elements FIRST though LAST (exclusive) from their
/users/yigonghu/Desktop/OS/pintos/src/lib/kernel/list.c
<http://www.gnu.org/software/gdb/documentation/>.
For help, type "help".
Type "apropos word" to search for commands related to "word"...
Reading symbols from kernel.o...done.
(gdb) target remote localhost:2430
Remote debugging using localhost:2430
The target architecture is assumed to be i8086
[f000:fff0]    0xffff0: ljmp   $0x3630,$0xf000e05b
0x0000fff0 in ?? ()
(gdb) b ../../lib/kernel/list.c:170
Breakpoint 1 at 0xc0029b39: file ../../lib/kernel/list.c, line 170.
(gdb) r
The "remote" target does not support "run".  Try "help target" or "continue".
(gdb) c
Continuing.
The target architecture is assumed to be i386
=> 0xc0029b39 <list_insert>:    push   %ebp

Breakpoint 1, list_insert (before=0xc0038220 <all_list+8>, elem=0xc000e020) at ../../lib/ker
nel/list.c:170
(gdb)
```

Debugging On Pintos

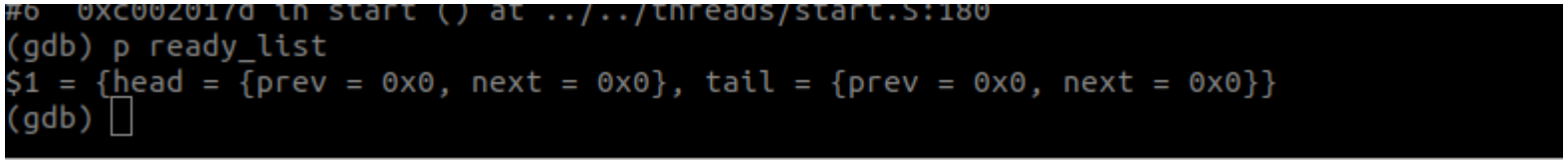
- Then we continue to run the program until it reaches the assertion error. We check the call stack of list_insert using `backtrace` (`bt`). We find that the thread_unblock function calls the insert_error.



```
220 }
221
222 /* Transitions a blocked thread T to the ready-to-run state.
223    This is an error if T is not blocked.  (Use thread_yield() to
224    make the running thread ready.)
225
226    This function does not preempt the running thread.  This can
227    be important: if the caller had disabled interrupts itself,
228    it may expect that it can atomically unblock a thread and
229    update other data. */
230 void
231 thread_unblock (struct thread *t)
232 {
233   enum intr_level old_level;
234
235   ASSERT (is_thread (t));
236
237   old_level = intr_disable ();
238   ASSERT (t->status == THREAD_BLOCKED);
239   list_push_back (&ready_list, &t->elem);
240   t->status = THREAD_READY;
241   intr_set_level (old_level);
242 }
243
244 /* Returns the name of the running thread. */
245 const char *
```

Debugging On Pintos

- We then check the value of ready_list and we find that the list is not initialized yet. Then we know the root cause is the uninitialized data.



Git

- Be familiar with daily commands
  - git clone
  - git commit
  - git push
  - ...
- A little more advanced
  - git checkout
  - git reset
  - ...
- Reference

  https://www.atlassian.com/git/tutorials

  https://www.katacoda.com/courses/git

Tips for using Git

- When using Git, you should:
  - Use .gitignore to avoid checking in unnecessary files (e.g., object files)
  - Write meaningful commit messages
  - Create branches for different exercises/features
  - Use pull requests to merge feature branch into master
    - Have teammates review the changes
  - Integrate your team's changes early and often
- When using Git, you should avoid:
  - Check in object files
  - Use git add -A command to commit all the files
    - Use git status to see modified & staged files
  - Use git push --force to force remote update

IDE settings: vim

- Vim is sufficient for daily development on servers
  - Recommended Vim plugins: NERDTree, YouCompleteMe, COC, fzf.vim
  - Recommended NeoVim plugins: COC.nvim
  - Plugin manager: vim-plug

- cscope is a developer's tool for browsing source code
  - Reference: http://tuxdiary.com/2012/04/03/code-browsing-using-ctags-and-cscope
  - Use cscope in Pintos: https://www.cs.jhu.edu/~huang/cs318/fall21/project/pintos_12.html#SEC170
    - `make cscope` in src/

IDE settings: vim

- "Inconvenient" to use at the beginning
  - Try to force yourself using them
  - Typing keys is in general much faster than mouse-click
  - Once you develop the habit, this set of tools will be your lifetime companions that boost programming productivity

IDE settings: VSCode

- VSCode is a GUI IDE by Microsoft
  - Free of charge
  - Easy to use for remote development via SSH

- Cache locally, sync automatically, execute remotely
  - Install “Remote Development Extension Pack” plugin
  - Connect to a remote host
  - Remote Development Reference: https://code.visualstudio.com/docs/remote/ssh

QEMU/bochs

- QEMU: faster

  pintos --qemu …

  make check SIMULATOR=--qemu

- Bochs: support reproducible mode

  pintos --bochs ...

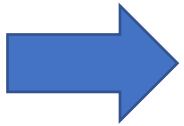
  make check SIMULATOR=--bochs

  (For reproducible mode, refer to http://bochs.sourceforge.net/doc/docbook/user/bochsrc.html, section 4.3.16)

  (Lab 1 will use bochs & Lab 2-4 will use QEMU. Change SIMULATOR for local test-only.)

Outline

- Administrivia
- Lab 0 overview
- Environment Setup
- Dev Tool
- Tips



Development Suggestions

- Bad coding habit
  - Divided the assignment into pieces
  - Each group member worked on his or her piece in isolation until just before the deadline
  - Reconvened to combine their code and submit

- Why is it bad?
  - Conflict with each other
  - Requiring lots of last-minute debugging

Development Suggestions

- Good coding habit
  - Integrating your team's changes early and often
  - Do incremental function testing
  - Using a source code control system such as Git
  - Read the compiler WARNINGs

- This is less likely to produce surprises
- These systems also make it possible to review changes and, when a change introduces a bug, drop back to working versions of code

Code Style

- Can your group member and TAs understand your code easily?
  - E.g.: GNU code style: http://www.gnu.org/prep/standards/
  - Limit source file & function lines
  - Following the naming convention
  - Adding meaningful comments
  - If you remove existing Pintos code, delete it from your source file entirely

- The style will be accounted for during grading
  - Bad code or code with messy styles will get points deducted even if it passes tests.

General Tips

- Think about design before you start coding
  - You can run through your design with the TAs or Professor if you are unsure
- Reserve enough time to write design doc (30% of score)
- Carefully read the project documentation pintos.pdf/.html
  - The appendix of the doc is particularly helpful
  - Many confusions come from not reading the documentation thoroughly

Reference

Project Website

https://yigonghu.github.io/EC440/fall25/projects/

Pintos

https://www.cs.jhu.edu/~huang/cs318/fall21/project/pintos_7.html

CGDB

https://cgdb.github.io/

Git

https://www.atlassian.com/git/tutorials

Recap C/Assemble Language

Reference for C:

https://en.wikibooks.org/wiki/C_Programming/Advanced_data_types

https://en.wikibooks.org/wiki/C_Programming/Pointers_and_arrays

Reference for assembly:

https://en.wikibooks.org/wiki/X86_Assembly/GAS_Syntax

https://wiki.osdev.org/Inline_Assembly

Thanks for attending! Happy hacking!