ECE 574 – Cluster Computing
Lecture 6

Vince Weaver

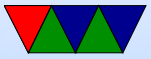
http://www.eece.maine.edu/~vweaver

vincent.weaver@maine.edu

17 September 2015

Announcements

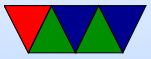
- Homework #3 will be posted soon



Haswell EP Info

Unpacked and set up new Haswell EP server: 16 cores (32 threads) 16GB RAM, 20MB cache each socket.
230 GFLOPS

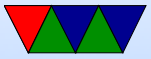
Roughly the same speed I measured on a 46-node, 184 core, Pentium D cluster I built when in grad school (though I didn't exhaustively test that cluster, it probably could have peaked higher)



Finding Detailed Stats on Modern CPUs

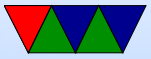
This is not always easy. There are papers, and manuals at the vendor websites, but these don't always present a clear picture.

One good resource is Agner Fogg's various optimization guides: `http://www.agner.org/optimize/`



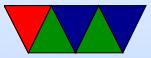
Example Problem for the Homework?

- Matrix multiply is typical, but boring
- What else can we use that's embarrassingly parallel, but interesting?



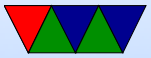
Convolution

- Specifically 2-D convolution
- Widely used in image processing
- Walk over every pixel in an image, convolving a matrix over it. The new value is based on some combination of the surrounding pixels.
- Usually a 3x3 grid, but can be larger



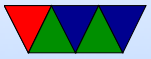
Common Convolution Matrices

- Identity = $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
- Blur = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ (need to normalize)
- Sharpen = $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$
- Emboss = $\begin{bmatrix} -2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$
- Sobel (edge detection) = $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$



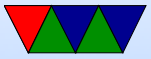
What the code will look like

```
for(x=1;x<width-1;x++) {
    for(y=1;y<height-1;y++) {
        sum=0;

        for(k=-1;k<2;k++) {
            for(l=-1;l<2;l++) {
                sum+=old[((y+l)*width)+(x+k)] *
                    filter[k+1][l+1];
            }
        }
        /* Normalize if necessary */

        /* Saturate if necessary */
        new[(y*width)+x]=sum;
    }
}
```



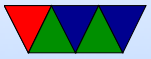
How to Optimize

- ROW vs Column Major? FORTRAN vs C? Comes down to using cache in an expected way.
- Loop order? Again, want to access in a way that keeps things in cache
- Loop unrolling? Avoids branch issues, etc.
- SIMD? Definitely a case where we could load all 4 channels and operate on them at once. Possibly multiple. A bit advanced for this class though.

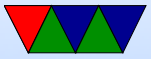


Can we Parallelize This?

- Yes we can!

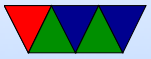


Parallel Programming!



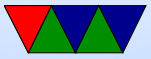
Processes – a Review

- Multiprogramming – multiple processes run at once
- Context switch – each process has own program counter saved and restored as well as other state (registers)
- OSes often have many things running, often in background.
  On Linux/UNIX sometimes called daemons
  Can use `top` or `ps` to view them.
- Creating new: on Unix its fork/exec, windows



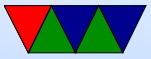
CreateProcess

- Children live in different address space, even though it is a copy of parent
- Process termination: what happens?
  Resources cleaned up. atexit routines run.
  How does it happen?
  `exit()` syscall (or return from main).
  Killed by a signal.
  Error
- Unix process hierarchy.



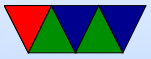
Parent and children, etc. not strictly possible to give your children away, although init inherits orphans

- Process control block.



Threads

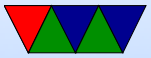
- Each process has one address space and single thread of control.
- It might be useful to have multiple threads share one address space
  GUI: interface thread and worker thread?
  Game: music thread, AI thread, display thread?
  Webserver: can handle incoming connections then pass serving to worker threads
  Why not just have one process that periodically switches?



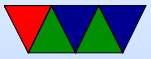
- Lightweight Process, multithreading

- Implementation:
  Each has its own PC
  Each has its own stack

- Why do it?
  shared variables, faster communication
  multiprocessors?
  mostly if does I/O that blocks, rest of threads can keep going
  allows overlapping compute and I/O

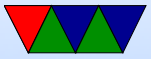


- Problems:
  What if both wait on same resource (both do a scanf from the keyboard?)
  On fork, do all threads get copied?
  What if thread closes file while another reading it?



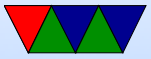
Thread Implementations

- Cause of many flamewars over the years



User-Level Threads (N:1 one process many threads)

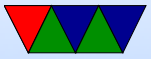
- Benefits
  - Kernel knows nothing about them. Can be implemented even if kernel has no support.
  - Each process has a thread table
  - When it sees it will block, it switches threads/PC in user space
  - Different from processes? When thread_yield() called it can switch without calling into the kernel (no slow



kernel context switch)
  - Can have own custom scheduling algorithm
  - Scale better, do not cause kernel structures to grow

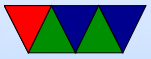
- Downsides
  - How to handle blocking? Can wrap things, but not easy. Also can't wrap a pagefault.
  - Co-operative, threads won't stop unless voluntarily give up.

    Can request periodic signal, but too high a rate is inefficient.

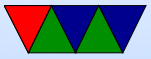


Kernel-Level Threads (1:1 process to thread)

- Benefits
  - Kernel tracks all threads in system
  - Handle blocking better
- Downsides
  - Thread control functions are syscalls
  - When yielding, might yield to another process rather than a thread

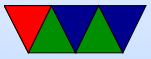


– Might be slower



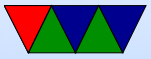
Hybrid (M:N)

- Can have kernel threads with user on top of it.
- Fast context switching, but can have odd problems like priority inversion.



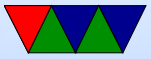
Linux

- Posix Threads
- Originally used only userspace implementations. GNU portable threads.
- LinuxThreads – use clone syscall, SIGUSR1 SIGUSR2 for communicating.
  Could not implement full POSIX threads, especially with signals. Replaced by NPTL
  Hard thread-local storage



Needed extra helper thread to handle signals
Problems, what happens if helper thread killed? Signals broken? 8192 thread limit? proc/top clutter up with processed, not clear they are subthreads

- NPTL – New POSIX Thread Library
  Kernel threads
  Clone. Add new futex system calls. Drepper and Molnar at RedHat
  Why kernel? Linux has very fast context switch compared to some OSes.
  Need new C library/ABI to handle location of thread-



local storage

On x86 the fs/gs segment used. Others need spare register.

Signal handling in kernel

Clone handles setting TID (thread ID)

exit_group() syscall added that ends all threads in process, exit() just ends thread.

exec() kills all threads before execing

Only main thread gets entry in proc