

KAWSE GROW and EXCITE Workshops

Summer 2013

High Performance Computing with Scratch

Learning Objectives

- See how computers play an important part in forecasting weather and climate change
- See how a single computer would be unable to provide the amount of power needed to create accurate forecasts and models in a reasonable amount of time
- See that we can easily break a weather forecast or climate model down to smaller parts and run those parts on multiple computers
- Understand how High Performance Computing helps us understand our world's weather and climate.
- Experience working with a program in Scratch designed to mimic some of the real-world gains and tradeoffs that come from using HPC.

Concepts, Terminology, People

- Supercomputer
- High Performance Computing (HPC)
- Resolution (Weather Maps)
- Threads and Processes
- Parallelism
- Resolution (granularity)

Resources

- NASA Advanced Supercomputing Division (NAS): <http://www.nas.nasa.gov>
- National Center for Atmospheric Research (NCAR): <http://ncar.ucar.edu/>
- MIT Scratch: <http://scratch.mit.edu/>

Lesson Setup Before Class

- (Optional) Make sure there is a guest account available on the computer labs
- Install the drivers for the Scratch PicoBoards on the computers (done 6/4/2013)
- Log on to the computers in the lab with either a user account or a guest account.
- Connect the Scratch PicoBoards and verify that they are working correctly.
<http://www.picocricket.com/picoboardsetupUSB.html> has good information
- Load the Scratch Starter file from the internet or other means (available at <http://people.cis.ksu.edu/~russfeld/presentations/kawse2013.sb>)

Lesson Plan

Presentation is online at <http://people.cis.ksu.edu/~russfeld/presentations/kawse2013.html>

1. [Intro] Today we are going to talk about using High Performance Computing to forecast the weather. I'm sure many of you have seen someone on TV talking about the weather, especially here in Kansas. Does anyone know how they come up with those weather forecasts? [Discuss]
2. [Quote] Here is a really great quote about forecasting weather: "The trouble with weather forecasting is that it's right too often for us to ignore it and wrong too often for us to rely on it" - Patrick L Young. It really is interesting that, even with all of our modern advances in science and technology, it is still extremely difficult to have an accurate weather forecast for more than a couple of days in the future.
3. [Map] This map comes from Weather.com, and shows the wind directions and wind speeds across the United States. Each arrow represents the direction of the wind, and the colors show how fast the wind is blowing. If we wanted to forecast what the winds would look like tomorrow, we would want to create a map very similar to this one.
4. [Processing Power] Unfortunately, to calculate tomorrows winds, it would take a TON of processing power to accurately predict the direction and speed represented by just a single arrow. [Draw line on the board] This line represents the amount of time it would take to do that calculation. [Draw another line directly after the first] And this is the next arrow. [You can continue drawing lines in succession to further illustrate the point]. As you can see, this could take a very long time to do.
5. [Independently] Thankfully, each arrow can be produced independently from the others. That means that we can be working on creating multiple arrows at the same time, like this: [draw a few parallel lines]. This idea is called "parallelism" and comes from the idea of working on multiple things in parallel, just like these lines.
6. [Computer] To do these calculations, we can use a computer. Today's computer can do billions of calculations each second, but it would still take a modern computer hours to compute a single arrow on that map.
7. [One thing] The other big problem is that most computers can only work on one thing at a time. [Refer to the series of lines on the board]
8. [Days] With just a single computer, it would take DAYS to calculate that map. By the time you are done calculating it, the data would already be too old to do any good.
9. [Resolution] So, we need to figure out a way to make it faster. One way to do that is to reduce the number of arrows we need to calculate. This can also be called the "resolution" of the calculation.
10. [High Res] This is a pretty high resolution image of the data...
11. [Med Res] but as we go to a lower resolution image...
12. [Low Res] ...it starts to look different. Do you think this image is going to be very useful if you are just worried about the weather in a particular city? [Discuss]
13. [Less Accurate] Unfortunately, by reducing the resolution, it can make your weather forecasts less accurate. We don't want that, do we?
14. [How?] So, my question to you is: how can we quickly and accurately forecast the

- weather using the technology we have available today? [Discuss, hopefully they talk about using multiple computers, but you may need to lead them in that direction]
15. [Computers] Maybe we could use something like this?
 16. [HPC] That's right! We can use High Performance Computing to spread the calculations across many different computers. [Refer to parallel lines] Since each of these lines now has a computer they can run on, we can do many things at the same time without slowing down the entire process.
 17. [Today] ... and today we are going to show you a little bit of how it works.

Scratch Activity: Weather Simulation

Break the students up into pairs/teams if needed, so that each group has a computer and a Scratch board (if they are used). Make sure they are able to open Scratch and can see the sample program we have provided. You can give a quick walkthrough of the program showing them how it works, but mainly focus on how they can change the variables in the program and/or using the slider on the Scratch board. The variables have the following valid ranges:

Threads: 1 through 12 inclusive

Resolution: 1 through 16 inclusive recommended, may go as high as 32 if desired

To start the program, have the students click the Green Flag, then adjust the slider to their desired spot. Press the button on the Scratch Board to run the simulation.

Have the students run several simulations (as time allows) with different combinations of threads and resolution, and try to graph the results against the time it took to complete the calculation. (It may be easiest to have them make a graph for each number of threads they use and plot resolution against total time or vice versa) See if they can find the number of threads for each resolution that gives the fastest overall time (it usually is around 4 to 8 threads in my testing).

Advanced Discussion: Why are the results very similar for 8 threads and 12 threads? (I believe it is due to the fact that the PCs we run the program on are using chips with 4 or 8 cores, so the 12 threads really only run on that many processing cores.)

18. [Uses of HPC] As you have seen, High Performance Computing can do great things when it comes to making our weather forecasts more accurate. However, that is just a small part of what it can really do.
19. [Watson] How many of you remember Watson, the computer that beat Jeopardy's two best champions?
20. [Watson Doctor] Well, they are now using it to learn everything there is to know about modern medicine, so that one day it can help doctors. Give it a list your symptoms, just like a Jeopardy question, and it will return a diagnosis quickly and accurately.
21. [Blue Gene] IBM has also created several High Performance Computing installations known as Blue Gene. They are used to do everything from sequencing the human genome...

22. [Cosmos] ... to working on simulations of the ENTIRE COSMOS, hoping to finally unravel the mysteries of the universe and understand how it can have very real effects on our climate.
23. [NASA] NASA is also a big user of High Performance Computing. They have several different supercomputers, including the Columbia seen here, that they can use...
24. [Rocket] ... to help blast their rocket science into the future by accurately simulating what it is like to send a rocket into space without actually needing a real rocket.
25. [Oceans] NASA is also working on simulating how ocean currents move around the world, which is vital to understand how a hot summer in Africa can affect the hurricane season in North America.
26. [NCAR] Another group doing lots of work with High Performance Computing is the National Center for Atmospheric Research, or NCAR.
27. [Wyoming] In October 2012, they opened their newest system in Wyoming called "Yellowstone" that is a monster. It is among the top 20 supercomputers in the world, and is capable of an astonishing 1.5 petaflops (or 1.5 quadrillion calculations per second).
28. [Globe] With that kind of power, you can quickly simulate wind conditions across the ENTIRE WORLD!.
29. [TwisterData] Finally, if you want to see a real world use of all of this data, check out TwisterData.com. They take all the latest information from all sorts of sources and provide it online free of charge to storm chasers and amateur meteorologists. When you want to stay safe during severe weather, access to the most accurate and recent data is crucial.
30. [Overview] Overall, you've seen lots of uses for High Performance Computing today. Can you think of any other things that might use High Performance Computing? [Discuss as time allows]

After Lesson

- Collect Scratch PicoBoards and verify all cables are present.
- Log out of the computers.