Parallel Computing Notebook
FELIX LIU
April 2017

Flipping Coins in Parallel
Model Website

Q: If each worker can flip one coin per time step, how many time steps does it take the serial worker to flip two coins?
A: 2
Q: How many time steps does it take two parallel workers to flip the two coins?
A: 1
Q: Set the work size to 16. How many time steps does it take the serial worker to flip the coins?
A: 16
Q: How many time steps does it take the parallel workers to flip the coins?
A: 8
Q: Set the number of parallel workers to 4. How many time steps does it take the parallel workers to flip the coins now?
A: 4
Q: Set the number of parallel workers to 8. How many time steps does it take the parallel workers to flip the coins now?
A: 2
Q: From what you've seen, what is one reason why it would be a good idea to use parallel workers instead of a serial worker?
A: Parallel workers take less time to complete the same task compared to a serial worker.
Q: Decrease the number of parallel workers to 2. Decrease the max time to 2. In 2 time steps, how many coins can be flipped by 2 parallel workers compared to one serial worker?
A: The parallel workers flip 4 coins while the serial worker flipped 2.
Q: Increase the number of parallel workers to 4. In 2 time steps, how many coins can be flipped by 4 parallel workers compared to one serial worker?
A: The parallel workers flip 8 coins while the serial worker flipped 2.
Q: Increase the number of parallel workers to 8. In 2 time steps, how many coins can be flipped by 8 parallel workers compared to one serial worker?
A: The parallel workers flip 16 while the serial worker flipped 2.
Q: From what you've seen, what is another reason why it would be a good idea to use parallel workers instead of a serial worker?

A: They complete more than a serial worker in the same amount of time.

Q: Increase the max time to 16 time steps. Decrease the number of parallel workers to 2. Decrease the max worker memory to 2 coins. If each worker can only hold 2 coins in memory, what is the maximum number of coins that can be flipped by 2 parallel workers compared to 1 serial worker?

A: Parallel : 4, Serial : 2

Q: Increase the number of parallel workers to 8. If each worker can only hold 2 coins in memory, what is the maximum number of coins that can be flipped by 8 parallel workers compared to 1 serial worker?

A: Parallel : 16, Serial: 2

Q: From what you've seen, what is another reason why it would be a good idea to use parallel workers instead of a serial worker?

A: Parallel workers will hold more memory than just one serial worker

Human Parallel Computer - Data Parallelism through Forest Fire Simulations

My number:19

Total number of students:22

My probability:19/22

Percentages:100, 100, 100, 100, 99.65

Iteration counts:17, 17, 17, 17, 17

Average percentage: 99.93%

Averge # of iterations: 17

Q: What were some of the **tasks** we did in this exercise? What were they, and who did them?

A: 1. Getting a sticky note with a number
2. Finding number of students
3. Calculating percentage
4. Running model 5 times
5. Recording data
6. Calculating averages
7. Returning data to Aaron

Q: What kinds of **data** did we work with in this exercise?

A: We worked with averages, probability and percentages.

Q: In which steps was there **communication** or **message passing** during this exercise (mark these steps)?

A: We communicated when we received the sticky notes and returned the sticky notes. We also communicated to the actual computer.

Q: In what ways could this exercise have been **optimized** so it could take less time?

A: Everyone could have entered in their own data in a shared graph.

Q: How could we have run this exercise using two instructors instead of one?

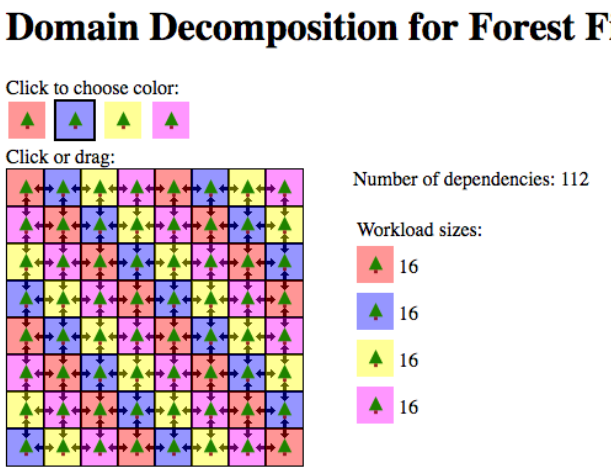
A: One instructor could have managed iterations while the other managed the percentage burned.

Q: In what ways did we simulate a **parallel computer** in this exercise?

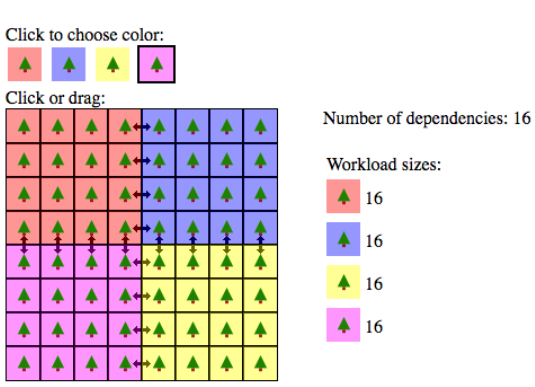
A: All students each ran the model with different specifications instead of one person running the data points in a series.

Domain Decomposition

Model Website



MAXIMUM NUMBER OF DEPENDENCIES



FOREST DIVIDED INTO FOUR EQUAL PIECES FOR EACH COLOR

Q: If we assumed each color is assigned to a researcher in a real forest, and each researcher is studying how a fire spreads through the forest, why do we call them **dependencies**? Why do we call them **workload sizes**?

A: They are called dependencies because a unit has to depend on others for information. They are called workload sizes because they only have so much memory they can use.

Q: If we assumed each color is assigned to a computer running a simulation for that part of the forest, why do we call them **dependencies**? Why do we call them **workload sizes**?

A: They are called dependencies because a researcher has to depend on others for information. They are called workload sizes because they only have so much memory they can use.

Q: Why would we want to minimize the dependencies?

A: We would want to minimize the dependencies because that would reduce communications.

Q: What are some reasons we might want to give more work to one of the colors/researchers/computers?

A: If one worker works better or faster than the others, it would be a good idea to give them more work.

Parallel Recipes

My serial recipe: How To Google Something

Materials:

1. Web browser
2. Computer
3. Keyboard
4. mouse

Instructions:

1. Press power button on computer to turn on
2. Put your hand on the mouse and move the cursor over a Google Chrome, Safari, or Firefox icon.
3. Click the button on the left side of the mouse
4. Once a new window is open, move the cursor over the long bar on the top of the window
5. Click the button on the left side of the mouse
6. Type "www.google.com" with the keyboard
7. Press the key that says "return" on your keyboard
8. There should be a long box in the center of the screen under the word "Google"
9. Move cursor on the box and click
10. Use the keyboard to type in what you want to Google.
11. Press the button on the keyboard that says "return"

Dependencies:

- Your computer has to be on in order to open a web browser.

My parallel recipe

Materials:

1. Web Browser
2. Computer
3. Keyboard
4. mouse

Instructions:

1. Billy presses power button on computer to turn on
2. Billy tells Jimbo when it has finished powering on and Jimbo puts his hand on the mouse and move the cursor over a google chrome, safari, or firefox icon and clicks
3. Billy tells Jimbo when the window is open, and Jimbo moves his cursor over the top bar on the window
4. Flash types in www.google.com with the keyboard
5. Flash presses the return button
6. Billy tells Jimbo when the screen is loaded and Jimbo clicks the search box
7. Jimbo tells Flash when he finishes clicking and Flash types in what they want to search
8. Flash presses enter

Q: In what ways was your **parallel** recipe different than your **serial** (non-parallel) recipe?

A: In my parallel recipe, there needed to be a lot of communication between members in order for success. Each member had different roles and had to work together to accomplish the task.

Q: In what ways was your **parallel** recipe the same as your **serial** (non-parallel) recipe?

A: Both had the same end product and followed a linear path.

Q: In what ways was your parallel recipe more efficient? In what ways was it less efficient?

A: No. It was less efficient because there needed to be clear communication between members.

Q: Did anything need to change about the resources/materials/ingredients/tools in your recipe when you went from serial to parallel?

A: No

Q: In what ways do you think this activity relates to computing and parallel computing?

A:It analyzes the differences between serial computing and parallel computing in completing a task.

Going Shopping

Solutions:

1. Each person goes into the supermarket and buys their own lunch
2. Aaron goes in and buys lunch for everybody
3. One person buys different ingredients to make a sandwich
4. Order food
5. Take each kid back to their house to retrieve the forgotten lunches

What if's:

1. You only had 100 dollars to spend
2. There were allergies
3. There are too many students to fit in the van
4. The van doesn't work
5. Nobody can drive
6. Get arrested for speeding
7. Supermarket is closed
8. The kids unite and plan a coup d'etat against Jim and Aaron

Q: Where is there inherent **parallelism** in your solutions?

A: Step 3: Each person has a specific thing to buy. They then use the ingredients to make a sandwich

Step 1: Each person buys their own lunch

Q: Where are there **dependencies** between tasks in your solution?

A: There could be allergies or the supermarket could be closed.

Q: Where is there **communication** in your solution?

A: The kids could communicate what they wanted to eat to Aaron or Jim.

Parallelism in Nature

1. Model link: https://www.agentcubesonline.com/project/40484
   Data: population in stores, timesteps
   Tasks: moving around, entering, exiting
   Parallelizable data: None
   Parallelizable tasks: Agents can move around at the same time.

Q: What patterns do you notice in the types of data and tasks that can be parallel?

A: They are things that can happen at the same time in one timestep.

Q: What patterns do you notice in the types of data and tasks that cannot be parallel?

A: Only one action can happen in one timestep.

Careers in High Performance Computing

Career: Software Developer

How HPC can be used in that career: A supercomputer needs well developed and designed software for it to work

Sources:

- http://www.cray.com/blog/the-software-imperative-in-high-performance-computing/

The World's "Fastest" Supercomputers

Q: When was the most recent Top500 list published?

A: November 2016

Q: What is the name of the fastest supercomputer in the world according to the most recent list?

A: Sunway TaihuLight – Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC

Q: Where is that supercomputer located?

A: Wuxi, China

Q: How many **cores** does it have?

A: 10,649,600

Q: How much **peak performance (RPEAK)** does it have?

A: 125,435.9

Q: How many of the Top500 sites in the top 10 are located in the United States?

A: 5

Q: If the **Blue Waters** supercomputer was capable of a **peak performance** of **13,000 TFLOP/S** when it came on-line in 2012, where would it be listed in the November 2012 list?

A: 3rd

Q: Why doesn't Blue Waters appear on that list?

A: They chose to not be included and opted out on appearing on the list. They did not think the list measured a supercomputer's power in the best way. Running linpack is only one small task a supercomputer can be used for..

Source: **https://www.hpcwire.com/2012/11/16/blue_waters_opts_out_of_top500/**

Q: What are **cores**?

A: A core is the basic computational unit of a CPU.

Q: What does **TFLOP/S** stand for?

A: Trillion Floating Point Operations per second.

Q: What does **Linpack** measure?

A: measure of a system's floating point computing power

Q: What would be some different ways to rank supercomputers?

A: You can rank them by power, performance per core, or Rmax.

LittleFe

Where the name comes from: Fe is comes from iron. Big iron is a slang term for supercomputing so littleFe means small iron because it's there's bigger computers.
Components:

- 6 motherboards
- 2 cores per CPU
- 1 hard drive
- Ethernet network
- RAM (memory)
- Every motherboard has a certain amount of memory
- Power supply
- Case
- cooling

Blue Waters demo

YouTube video

Q: What are the advantages to using a remote supercomputer as compared to a local supercomputer like LittleFe?
A: You can visualize things. It is a much powerful computer as well. You don't have to travel. Anyone can use Blue Waters.
Q: What are the disadvantages?
A: It isn't portable and is extremely time consuming and costly to build a supercomputer. You also need an account to use it.

Parallel Computing: Terminology and Examples

Slides

Shopping for Your Own Supercomputer

- PART: Motherboard (12)
- COST: $59.99 per motherboard (Total: 719.88)
- LINK: https://www.newegg.com/Product/Product.aspx?Item=N82E16813128565
- PART: Memory
- COST: 34.99
- LINK:https://www.newegg.com/Product/Product.aspx?Item=N82E16820231277&cm_re=memory_ddr3-_-20-231-277-_-Product
- PART: Hard Drive

- COST 16.99
- LINK: https://www.newegg.com/Product/Product.aspx?Item=9SIA5AD4HV9323
- PART: Ethernet Cables (12)
- COST: 3.88 per cable (Total: 46.56)
- LINK:https://www.newegg.com/Product/Product.aspx?Item=N82E16812189016&cm_re=ethernet_cable-_-12-189-016-_-Product
- PART: ROUTER
- COST: 35.49
- LINK:https://www.newegg.com/Product/Product.aspx?Item=N82E16833124190&cm_re=router-_-33-124-190-_-Product
- FAN
- COST: 39
- LINK: https://www.newegg.com/Product/Product.aspx?item=N82E16835103052
- CPU
- $85
- LINK: https://www.newegg.com/Product/Product.aspx?Item=N82E16819113286&ignorebbr=1

TOTAL: $977.91

Q: What are the most important parts of a supercomputer?

A: Motherboard, CPU

Q: What assumptions did you make when you did your shopping? Which of these assumptions were false?

A: I assumed that the best rated items that were also relatively cheap were the best buy. I feel like this assumption is true because if customers are both satisfied with the product and can buy it at a lesser price, then it's a really good buy.