

Parallel Computing Notebook

Ameya Rao

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: It shortens the time needed to flip all the coins.

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: 4:2

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: 8:2

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: 16:2

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: You can get more work done in a shorter 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: 4:2

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: 16:2

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: You get more done with more workers.

Human Parallel Computer - Data Parallelism through Forest Fire Simulations

My number:

Total number of students:

My probability:

(I wasn't in class this day, so I wasn't able to participate in the activity, but I experimented with different probabilities, using the Forest fire model)

Percentages:

Iteration counts:

Average percentage:

Average # of iterations:

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

A: Each person was given a number as well as a percentage.

Each person had to run their model and collect the data to give back to Aaron.

Aaron would then enter the data in a graph.

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

A: Probability, iteration, and the number.

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

A: All the steps used communications, between Aaron and the "parallel worker."

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

A: If each person entered their individual data it would have taken less time, because in this model Aaron had to enter all the data.

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

A: The two instructors could have entered data together, taking less time.

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

A: Everyone ran their own model, collecting data for one program.

Domain Decomposition

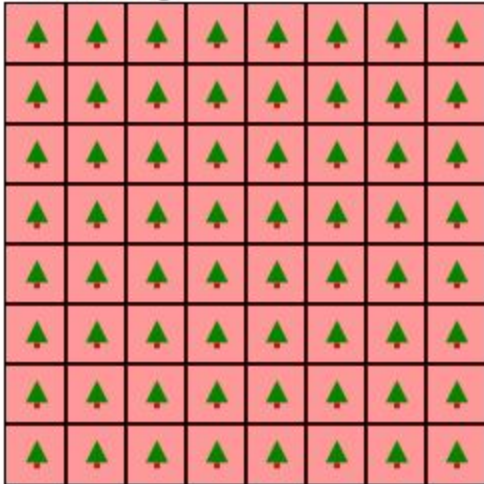
[Model Website](#)

Image:

Click to choose color:

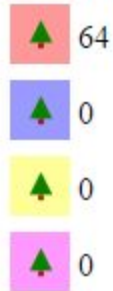


Click or drag:



Number of dependencies: 0

Workload sizes:



This is with no dependencies

Click to choose color:



Click or drag:



Number of dependencies: 16

Workload sizes:

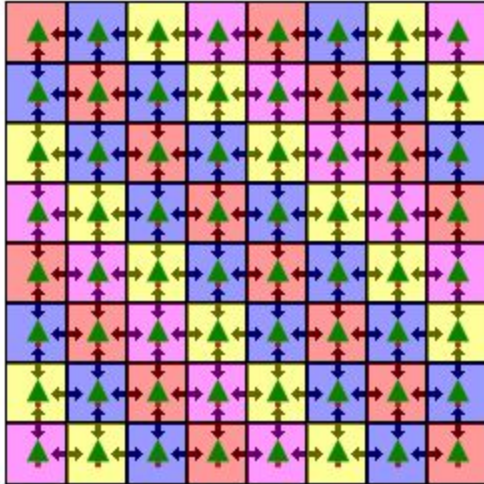


All the colors working in their own area so the dependencies are less.

Click to choose color:

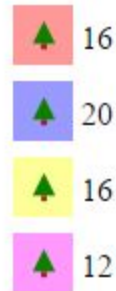


Click or drag:



Number of dependencies: 112

Workload sizes:



No tree is ever touching a tree of the same color, so this shows maximum dependencies.

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: The dependencies show that the researchers have to work with the other researchers to know what information the others have. If one tree is on fire around your tree, you are dependent on the other researcher to tell you if the tree is on fire. The workload size shows which researcher is in charge of which tree.

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: The dependencies show where the colors interact with each other and are dependent on what's going on in a different color. The workload sizes shows how many trees each color is in charge of.

Q: Why would we want to minimize the dependencies?

A: This makes the work more efficient because each color can work more independently and not have to wait for information from another color.

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

A: The other color might be working faster or their area is more secluded.

Parallel Recipes

My serial recipe

Materials:

1. Cheese
2. Bread

3. Pan
4. Stove
5. Butter
6. Knife
7. Cutting Board
8. Fridge
9. Drawer
10. Spatula
11. Plate
12. Cabinet

Instructions:

1. Open the cabinet A
2. Take out a plate A
3. Close the cabinet A
4. Open the Fridge B
5. Take out the Cheese B
6. Take out the Bread B
7. Take out the butter B
8. Close the fridge. B
9. Open the bread packet B
10. Take out two slices of bread B
11. Put them on the plate B
12. Open the cheese slice C
13. Place on top of one bread C
14. Open the cabinet D
15. Take out the pan D
16. Close the cabinet D
17. Place the pan on the stove D
18. Turn on the stove D
19. Place one slice of bread on top of the other B
20. Open the butter A
21. Spread on the pan A
22. Take spatula and knife out of drawer A
23. Close drawer A
24. Use spatula to put sandwich on stove A
25. Wait for 2 mins A
26. Flip the sandwich A
27. Wait 1 minute A
28. Flip the sandwich over onto the Cutting board. A
29. Turn off the stove. B
30. Use the knife to slice sandwich. A
31. Place sandwich halves on plate. A
32. Close bread packet D

33. Open fridge D
34. Put cheese, bread, and butter back in D
35. Put pan and knife in the sink. E

Dependencies:

- The pan must be on the stove before putting butter and turning it on.
- The sandwich cannot be sliced until it is fully cooked
- Can't open the cheese unless you have bread to put it on

My parallel recipe

Materials:

1. Same as before

Instructions:

1. Letters indicate which parallel worker does which step.
2. There are 5 total parallel workers.

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

A: There isn't any difference between the two recipes.

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

A: It had all the same steps.

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

A: The task was completed faster with more workers. It might be inefficient because you have a lot of people working in the same place.

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

A: No I didn't.

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

A: It shows the positives and negatives of having more people working on one assignment at once.

Going Shopping

Solutions:

1. Take all the kids in the Van to the Supermarket. Each one buys the food they want.
2. Jim drives to the supermarket and gets the food for everyone while Aaron teaches class.
3. Aaron drives to the supermarket and gets the food while Jim throws a party with the kids.
4. Kids call their parents and ask them to bring them lunch.
5. Everyone gets in the Van and they stop at everyone's house to pick up lunch.
6. Everyone gets in the Van and goes and buys pizza.
7. Half the kids go to the Supermarket and buy food for the other half as well.

What if's:

1. Kids have allergies.
2. Supermarket is closed.
3. Aaron and Jim can't drive.
4. Too many kids to fit in the van.

5. There isn't enough money.
6. They don't have the keys to the van.
7. They don't know how to get to the supermarket.
8. They are locked inside of Shodor.
9. Get pulled over for speeding.
10. Kids decide to go crazy and run around everywhere and refuse to get in the Van.
11. Kids get on a sugar high from eating too many brownies and rebel.

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

A: Jim teaches and Aaron buys food with the kids. Solving two tasks with two people. Each kid goes and buys their own food. Each parent comes and brings in their own food. Split the group in half, each kid is responsible for another too.

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

A: The supermarket needs to be open and have food. You need to have keys to drive the van. The van needs to have enough seats for all the kids. The food must be sufficient for all the kids, taking into consideration allergies and dietary preferences.

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

A: The kids must cooperate with Aaron and Jim. You have to contact the parents. You have to communicate with the cashier to buy the food.

Parallelism in Nature

1. Model link: <http://www.shodor.org/interactivate/activities/Chairs/>
 Data: Number of chairs needed, length and width.
 Tasks: Guessing how many chairs you need for a table
 Parallelizable data: None
 Parallelizable tasks: None
2. Model link: <https://www.agentcubesonline.com/AOO/Ristretto3D.html?nid=403877&mode=play>
 Data: Number of organisms, location, age
 Tasks: Movement of the fish, predators, and fishnet
 Parallelizable data: The program calculates the number of organisms simultaneously for each type
 Parallelizable tasks: All of the fish are moving and/or being eaten at the same time.

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

A: They aren't dependent on another factor at that time, so they can be calculated independently.

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

A: They are dependent on another factor occurring before the task is completed or data is collected. The waiting period allows for inefficiency with using parallel.

Careers in High Performance Computing

Career: Physicist

How HPC can be used in that career: Simulate the string networks for the investigation of string theory and analysis, Algorithm and Code development, Nonlinear gyrokinetic simulations, development of

increasingly detailed, reliable, and predictive models of complex systems and problems. Supernova Simulation, analysis of cosmic microwave background data.

Sources:

- <https://sites.tufts.edu/physics/shop/hpc-computer-cluster/>
- <http://www2.physics.uiowa.edu/~ghowes/research/numsim.html>
- <http://rspa.royalsocietypublishing.org/content/467/2131/1880>
- <http://crd-legacy.lbl.gov/~dhbailey/dhbtalks/dhb-desy.pdf>

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

Q: Where is that supercomputer located?

A: National Supercomputing Center in 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 TFLOP/S

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: The Blue Waters computer doesn't appear on the list because they did not like the way the list was calculated. They wanted to use real life applications.

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

Q: What are **cores**?

A: A central processing unit

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

A: Teraflops, 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: cores, rmax, rpeak, power, how many projects, types of problems being solved, efficiency (problems per core), energy efficiency,

LittleFe

- Cluster computer

Where the name comes from: "Big Iron" is a big supercomputer. So Fe is the chemical symbol for Iron, therefore Little Iron because it is a miniature supercomputer.

Components:

- Motherboards (6)
- 2 cores per CPU (12 total)
- Hard Drive (1)
- Network (Ethernet)
- RAM (Memory- 1,2, or 4 GB)
- Power Supply
- Cooling System (Fan, Heat Sink)
- Case
- Ethernet Router

Blue Waters demo

[YouTube video](#)

Q: What are the advantages to using a remote supercomputer as compared to a local supercomputer like LittleFe?

A: A remote supercomputer is much larger so it is capable of doing a lot more. More people can use BlueWaters at once and it will still run efficiently. It also has a lot more processing power, and you can connect without being near it.

Q: What are the disadvantages?

A: Since LittleFe is smaller, it is portable and you can take it with you. With Bluewaters, it has to stay in one place. You can see the visualizations. BlueWaters needs a lot of cybersecurity and takes up a lot of energy.

Parallel Computing: Terminology and Examples

[Slides](#)

Shopping for Your Own Supercomputer

- Newegg
- Amazon

- **Part:** Motherboard (x5)

Cost: \$38.00 x 5 = \$190.00

Link: [Motherboard Link](#)

- **Part:** Harddrive (x5)

Cost: \$7.99 x 5 = \$39.95

Link: [Hardrive Link](#)

- **Part:** Fan (Cooling System)

Cost: \$19.06 (Pack of Five)

Link: [Fan Link](#)

- **Part:** Memory RAM

Cost: \$95.99 (pack of 3- 4GB each) and \$71.81 (Pack of 2- 4GB each)

Link: [3 pack Link](#), [2 pack Link](#)

- **Part:** Ethernet Cable
Cost: \$9.99 (Pack of 5)
Link: [Ethernet Cable Link](#)
- **Part:** Ethernet Router
Cost: \$9.95
Link: [Ethernet Router Link](#)

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

A: The most important parts of a supercomputer are probably the motherboard and the ethernet.

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

A: I assumed that the more expensive motherboard would be better, but when I looked at the reviews, some of the cheaper ones are better.