Parallel Computing Notebook
Ritvik Bhagawatula
April 2017

Flipping Coins in Parallel
Model Website

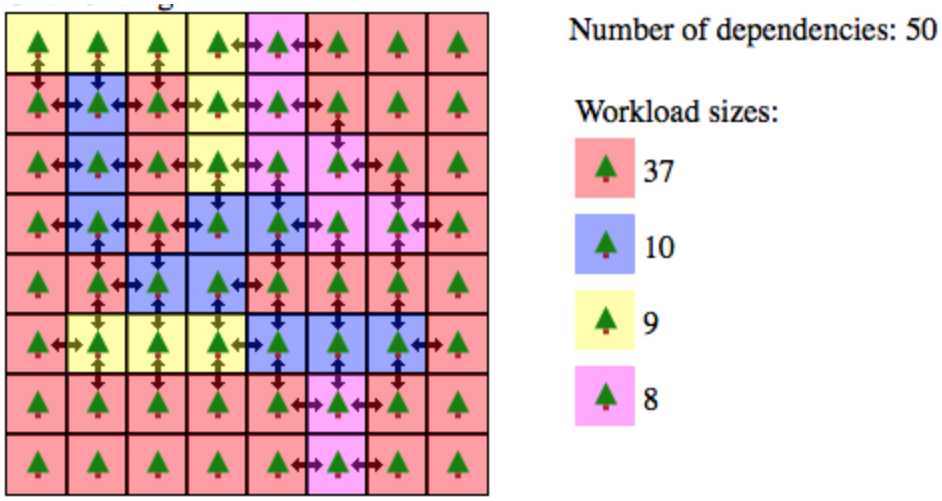
Q: If each worker can flip one coin per time step, how many time steps does it take the serial worker to flip two coins?
A: 2 Time Steps
Q: How many time steps does it take two parallel workers to flip the two coins?
A: 1 Time Step
Q: Set the work size to 16. How many time steps does it take the serial worker to flip the coins?
A: 16 Time Steps
Q: How many time steps does it take the parallel workers to flip the coins?
A: 8 Time Steps
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 Time Steps
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 Time Steps
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: Work gets done faster
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 Coins
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 Coins
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 Coins
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 can get more amount of work done without wasting 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 Coins

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 Coins
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 can get more done even with limited memory

Human Parallel Computer - Data Parallelism through Forest Fire Simulations
My number: 18
Total number of students: 22
My probability: 81.81%
Percentages: 100%, 99.65%, 98.26%, 99.65%, 100%
Iteration counts: 17, 17, 17, 17, 17
Average percentage: 99.512%
Averge # of iterations: 17
Q: What were some of the **tasks** we did in this exercise? What were they, and who did them?
A: Found averages and completed fire model
Q: What kinds of **data** did we work with in this exercise?
A: ID, Probability, Burn%
Q: In which steps was there **communication** or **message passing** during this exercise (mark these steps)?
A: Writing info on sticky notes and giving it to the instructor
Q: In what ways could this exercise have been **optimized** so it could take less time?
A: There could have been more instructors overseeing the process
Q: How could we have run this exercise using two instructors instead of one?
A: One instructor works with half the people while the other works with the other half.
Q: In what ways did we simulate a **parallel computer** in this exercise?
Instead of the instructor doing everything we worked as well

Domain Decomposition
Model Website

Image:



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: Because each cell is based on the previous one.
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: Because the fire is spreading throughout and it is getting weaker as it spreads
Q: Why would we want to minimize the dependencies?
A: We want to minimize the linkages between each cell and equally split the work
Q: What are some reasons we might want to give more work to one of the colors/researchers/computers?
A: To equally split the work between computers

Parallel Recipes

My serial recipe

Materials:

1) Car
2) Driver Permit
3) License
4) Birth Certificate
5) SSN

Instructions:

1. Attend the class
2. Drive with a certified instructor
3. Gather all the necessary documents
4. Go to DMV
5. Take the test and pass
6. Obtain permit

7. Drive for an year and for at least 60 hours
8. Go back to DMV to get provisional license
9. Get full license when 18
10. Register for a driving class

Dependencies:

- Parents
- Car Gasoline
- Weather
- Time
- Money

My parallel recipe

Materials:

1) Car
2) Driver Permit
3) License
4) Birth Certificate
5) SSN

Instructions:

1. Attend the class & Drive with a certified instructor
2. Gather all the necessary documents
3. Go to DMV
4. Take the test and pass
5. Obtain permit
6. Drive for an year and for at least 60 hours
7. Go back to DMV to get provisional license
8. Get full license when 18

Q: In what ways was your **parallel** recipe different than your **serial** (non-parallel) recipe?
A: With the parallel recipe the work is being done simultaneously
Q: In what ways was your **parallel** recipe the same as your **serial** (non-parallel) recipe?
A:Certain steps can only be done by one person and those did not change
Q: In what ways was your parallel recipe more efficient? In what ways was it less efficient?
A: There are two people working on the parallel recipe
Q: Did anything need to change about the resources/materials/ingredients/tools in your recipe when you went from serial to parallel?
A: No, because materials stay the same throughout
Q: In what ways do you think this activity relates to computing and parallel computing?
A: I think it's the same because some things only one computer can do.

Going Shopping

Solutions:

1. 1 person goes and buys all the food
2. 12 Kids go and bring food back

What if's:

1. What if they don't have enough money
2. What if the kids take over
3. What if the van breaks down

Q: Where is there inherent **parallelism** in your solutions?
A: One adult is watching over the kids while the other is going out to get the food so they are splitting their work.
Q: Where are there **dependencies** between tasks in your solution?
A: There are dependencies because the Kids can only eat their food once an adult goes and gets it
Q: Where is there **communication** in your solution?
A: The kids all have to tell the adult what they want so that would be communication

Parallelism in Nature

1. Model link: http://www.shodor.org/interactivate/activities/RabbitsAndWolves/
   Data: How much of the rabbits, wolves, and grass is left, where are they located. How much food they have eaten
   Tasks: The eating going on.
   Parallelizable data: One half of the wolf pack could take care of a half of the forest while the other half takes care of the other half of the forest
   Parallelizable tasks: The rabbits are moving around while the grass is growing.

Q: What patterns do you notice in the types of data and tasks that can be parallel?
A: They are types of data that do not have dependencies between each other.
Q: What patterns do you notice in the types of data and tasks that cannot be parallel?
A: The being eaten and growing because it needs to regrow.

Careers in High Performance Computing

Career: Computer Engineering
How HPC can be used in that career:
Sources: HPC can be used by a computer engineer in order to split complex computer processes into easier processes

- http://www.explainthatstuff.com/how-supercomputers-work.html

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: National Supercomputing center in Wuxi
Q: Where is that supercomputer located?
A: 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: Because they don't want to spend their time solving matrices to get on that list
Source:
Q: What are **cores**?
A: Cores are the cpu's which do the processing
Q: What does **TFLOP/S** stand for?
A: Teraflops per second
Q: What does **Linpack** measure?
A: A library for performing linear math operations on computers.
Q: What would be some different ways to rank supercomputers?
A: How many complex processes has this computer processed

LittleFe

Where the name comes from: The name means little iron
Components:

- Motherboards (x6)
- Cores (2 per CPU)
- Hard Drives (x1)
- Networks (Ethernet)
- Ram (Memory)

Blue Waters demo

YouTube video

Q: What are the advantages to using a remote supercomputer as compared to a local supercomputer like LittleFe?
A: With blue waters anyone from around the world can use the computer
Q: What are the disadvantages?
A: Because it takes a lot of space and energy to host it and with littlefe you can connect to it very easily.

Parallel Computing: Terminology and Examples
Slides

Shopping for Your Own Supercomputer

- Part: Motherboard + CPU
  Cost: $469.99
  Link:
  https://www.amazon.com/Z10PE-D16-WS-LGA2011-v3-CrossFireX-Motherboard/dp/B00QC5DZEU/ref=sr_1_7?ie=UTF8&qid=1495656613&sr=8-7&keywords=motherboard+%2B+cpu

- Part: Hardrive
  Cost: $119.19
- Link:
  https://www.amazon.com/Seagate-Expansion-Desktop-External-STEB5000100/dp/B00TKFEEBW/ref=sr_1_2?s=electronics&ie=UTF8&qid=1493488270&sr=1-2&keywords=hard drives

- Part: Memory
  Cost: $594
  Link:
  https://www.amazon.com/Kingston-Technology-HyperX-HX430C15PB3K4-64/dp/B01GCWQ72O/ref=sr_1_2?s=pc&ie=UTF8&qid=1493487779&sr=1-2&keywords=memory&refinements=p_n_feature_five_browse-bin%3A9559992011

- Part: Power Supply
  Cost: $80
- Link:
  https://www.amazon.com/EVGA-Modular-Warranty-Supply-210-GQ-0750-V1/dp/B017HA3RGE/ref=sr_1_5?s=electronics&ie=UTF8&qid=1493486121&sr=1-5&keywords=power+supply

- Part: Ethernet Switch
  Cost: $17.95
  Link:
  https://www.amazon.com/NETGEAR-ProSAFE-FS105NA-5-Port-Ethernet/dp/B00002EQCW/ref=sr_1_6?s=electronics&ie=UTF8&qid=1493486100&sr=1-6&keywords=ethernet+switch

- Part: Ethernet
  Cost: $6.29
  Link::

https://www.amazon.com/AmazonBasics-RJ45-Cat-6-Ethernet-Patch-Cable-10-Feet-3-Meters/dp/B00N2VIALK/ref=sr_1_3?s=pc&ie=UTF8&qid=1493486057&sr=1-3&keywords=ethernet

Q: What are the most important parts of a supercomputer?
A: The cores, storage equipment, and network connectivity are most important.
Q: What assumptions did you make when you did your shopping? Which of these assumptions were false?
A: The assumptions I made were that more is always better, but you only need enough storage and processing power as your tasks will take up.

Daqri demo

Q: What are some of the ways you can envision augmented reality being used for science?
A: I can see augmented reality being used in science with a surgeon who can send a small probe into the body and operate with a VR headset for maximum scope and clarity.
Q: What specific scientific examples can you think of that would benefit from augmented reality?
A: Surgeon, Laboratory.