When to Use OpenMP vs. MPI vs. Hybrid OpenMP+MPI

Goal

- Here, we are going to look at the higher level concepts associated with programming for shared memory and distributed memory combined
- Look at when to use a hybrid OpenMP+MPI model vs. an OpenMP model vs. an MPI model
- Refresh: What is shared memory? What are examples of when it's useful?
- Refresh: What is distributed memory? What are examples of when it's useful?

Introduction

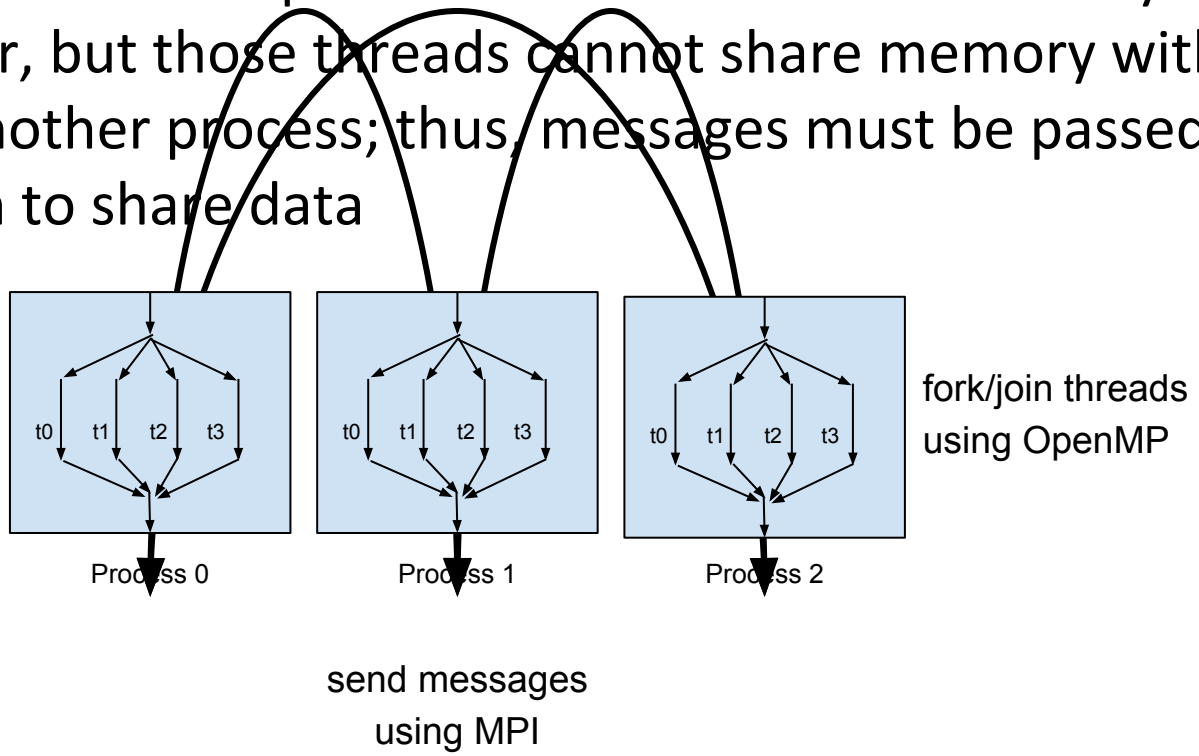
- Hybrid parallelism maximizes of the strengths of distributed memory and shared memory at the same time
- One possible combination of programming models for this is OpenMP+MPI

Why Hybrid Parallelism over Other Types?

- Works well for complex problems that can split data up across distributed memory and do a large number of complex computations
- Works for problems that would take too long to run or wouldn't be feasible without using the maximum amount of parallelism possible
- Works for problems that can be divided across distributed memory and those disjoint parts (with some communication) can be split across multiple threads

How Does it Work?

- Multiple MPI processes exist and each process forks into multiple OpenMP threads
- Threads of one process are able to share memory with each other, but those threads cannot share memory with threads of another process; thus, messages must be passed back and forth to share data



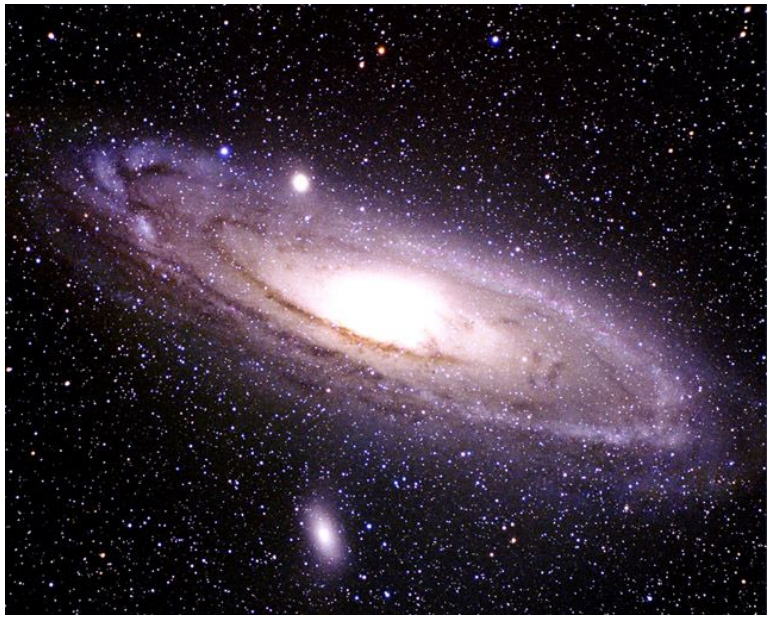
Exercises

Exercise #1: When to Use Hybrid?

- Now that we understand what hybrid parallelism is and why you might use it, let's look at some examples and see if we can figure out what type of parallelism would be the most useful for each: shared memory, distributed memory, or hybrid shared+distributed memory
- Note that there are many times where any type of parallelism would provide some type of speedup, but the goal is to find the right type of parallelism for the specific problem you're looking at

Exercise #1, cont.

- An astronomical model where we want to analyze 16 disjoint galaxies. Calculations will be run on every individual star in each galaxy



OpenMP, MPI, or hybrid?

WHY?

Exercise #1, cont.

- We want to calculate the photosynthetic rate of each individual leaf of 10,000 different plants. The photosynthetic rate of one plate does not affect the photosynthetic rate of another, but the photosynthetic rate of one leaf of a plant can affect the photosynthetic rate of another leaf on the same plant

OpenMP, MPI, or hybrid?

WHY?

Exercise #1, cont.

- We need to analyze the amount of traffic that goes through a neighborhood with 4 intersections to determine whether or not putting traffic lights in would be helpful

OpenMP, MPI, or hybrid?

WHY?

Exercise #1, cont.

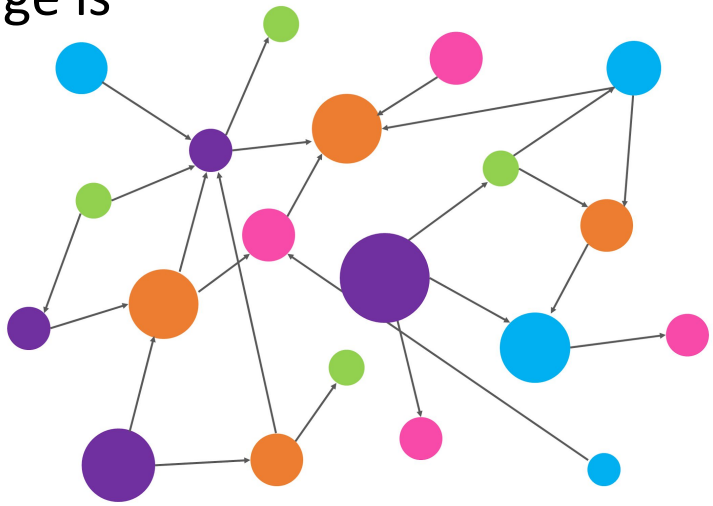
- We will perform a scalability student on an image processing algorithm to see how well the algorithm scales with really large datasets. The example dataset you will run your strong and weak scalability studies on is 4 TB in size

OpenMP, MPI, or hybrid?
WHY?

Exercise #1, cont.

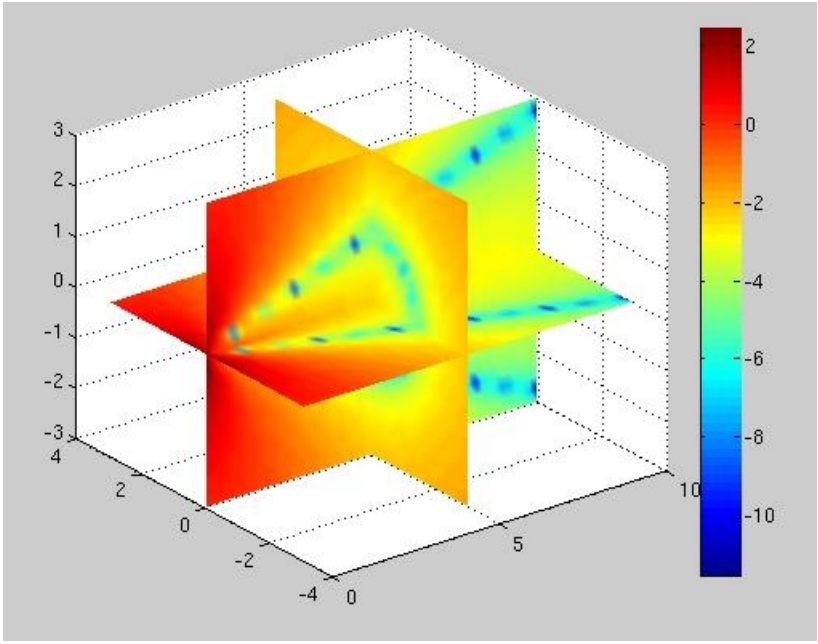
- The goal is to build a graph that represents a photograph of a landscape as seen from up above (image that it is a picture taken from an airplane). The nodes of the graph will represent different sections of the landscape, but each pixel of the graph requires information about other pixels so that it can determine whether or not the pixels are in the same section of the graph that is being generated. Your image is 512x512 pixels

OpenMP, MPI, or hybrid?
WHY?



Exercise #1, cont.

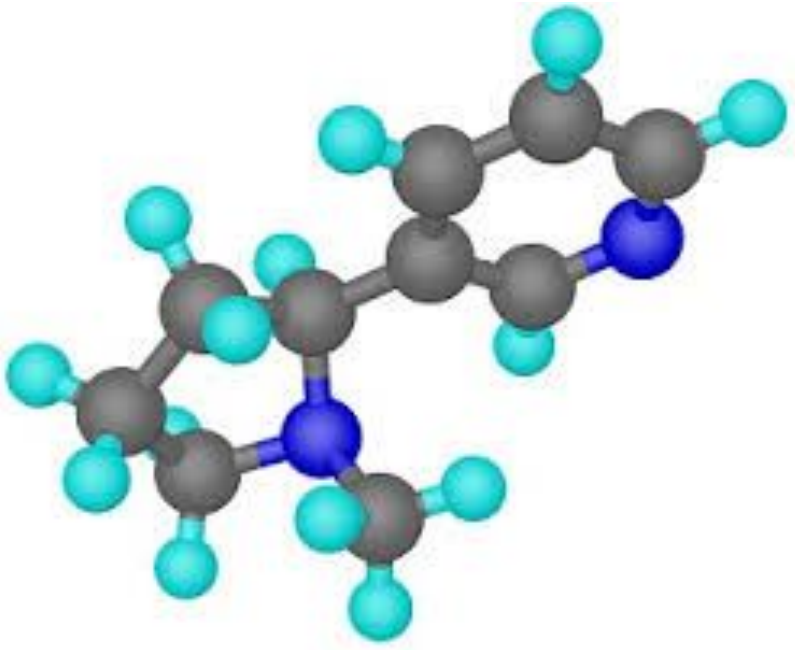
- The visualization we want to build contains 27 billion total elements laid out as a cube. There are 3072 files that will be read into the visualization software. Each file is a sub-grid of the full cube that contains 192x129x256 cells



OpenMP, MPI, or hybrid?
WHY?

Exercise #1, cont.

- We want to run 8 completely different chemical reactions and from the data gathered, we can the build the output molecular model. Two additional pieces of information are gathered from each chemical reaction: the final amount of the reactant and the final amount of the product



OpenMP, MPI, or hybrid?
WHY?

Exercise #2: Running Applications with OpenMP, MPI, and OpenMP+MPI

- This example involves running numerous different applications with OpenMP, with MPI, and with hybrid OpenMP+MPI to see which version works the best for each application
- Analyze the results from the combination of version and application as well as the scalability of each application with each version
- Instructions are in Exercise Instructions document