WHAT IS XSEDE…
AND SO CAN YOU!

Josh Faber

WHAT COMPUTERS DO WELL

- Math
- Graphics
- Match-3 games

WHAT COMPUTERS DON'T DO AS WELL

- Transfer data
- Move data around
- Tell jokes

ORDERS OF MAGNITUDE

- These days, anything on the order of a million operations is “fast” -- the time it takes will be around a second, and depend on external factors

- If you need to write a computer code that will run in a matter of seconds, consider just getting it done, and refining it later

ORDERS OF MAGNITUDE, II

- If your code needs a billion (10^9) operations, ir may not be fast anymore.

- Now you need to think about how long the code will take, and consider how to program it efficiently
  - Reuse math?
  - Order of calculations/dependencies
  - Memory access/proximity

WHY A BILLION

- In numerical work, we talk about the order of a method:
- A method of order "n" satisfies the property

$$\text{Error} \propto (\text{Scale})^n$$

- In practice, the amount of work goes like 1/scale
- That means:
  - First order methods <--> Twice the work, half the error
  - Fourth-order methods <--> Twice the work, 1/16 error

THE WORLD WE LIVE IN

- In 1-d, if you need a billion points to resolve something, you are probably doing it wrong
- In 2-d, you are still probably doing it wrong
  - Cameras are now in the muli-megapixel range, still generally smaller than $10{,}000^2$
- In 3-d, you can get to a billion easily: $(10^3)^3$
- 1,000 points/steps is not a lot in the numerical sense

GRAPHICS

- 1 million pixels is not a lot, 1 billion pixels would be
- The number of pixels in a movie is potentially gigantic
  - Several million pixels x 24 frames/sec * ??? seconds
- Potentially a billion RGB values every few seconds

PARALLEL COMPUTING

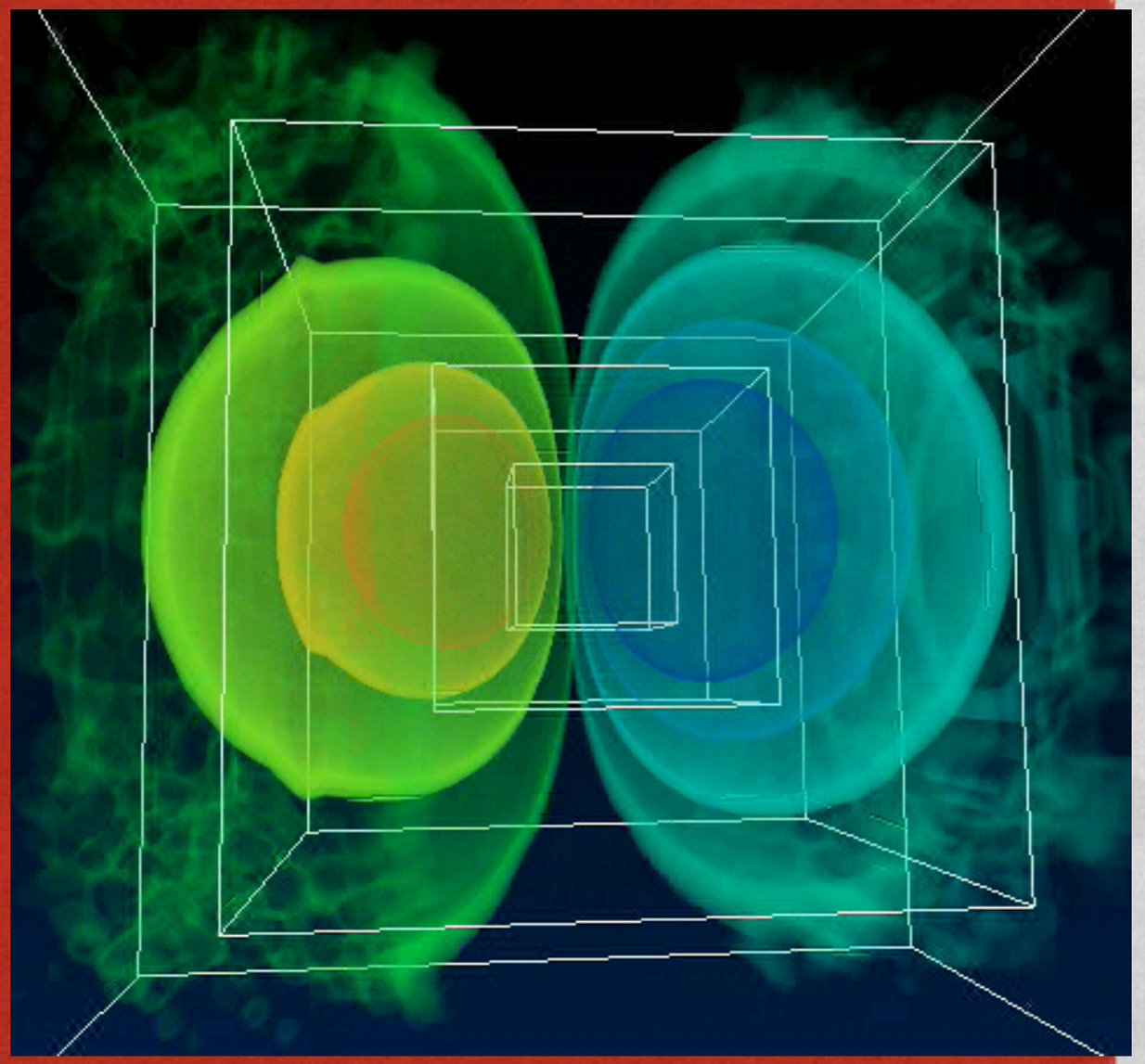
- We can use the power of many computers to make work go faster
  - “Strong” scaling - Use more computers to make same problem go faster
  - “Weak” scaling - Use more computers to make bigger problems take the same amount of time
- The secret is to let the computers work on math, and minimize the communication

TRIVIAL PARALLELIZATION

- Frame-by-frame video generation, assuming data files are staged correctly
- Monte-Carlo methods (not Markov chain!)
- To simulate binary evolution, we pick binary star masses, separations, supernova explosion parameters, etc., from known probability distributions
- At the end, tally all the results from different machines

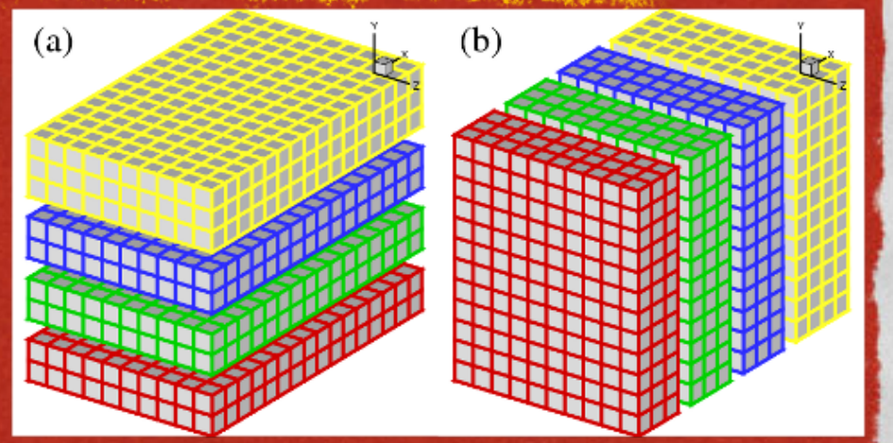
NOT-AS-TRIVIAL PARALLELIZATION

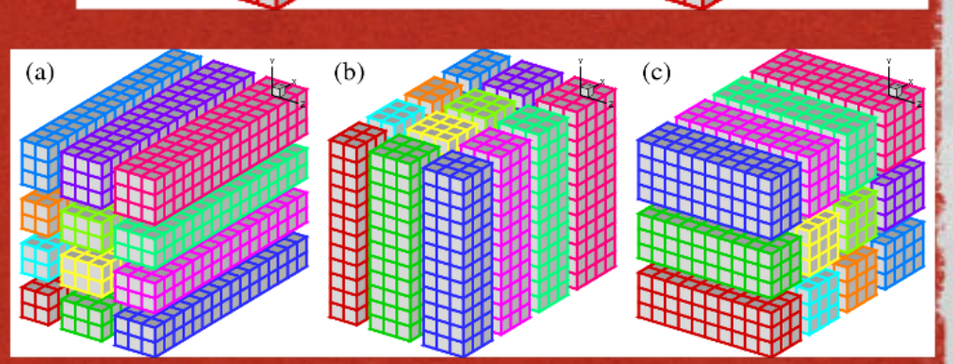
- 3-d grids with stuff
- Markov chain Monte Carlo with many-dimensional phase spaces
- Anything with many-dimensional phase spaces
- 3-d elliptic equations (linear algebra w/ giant matrices)

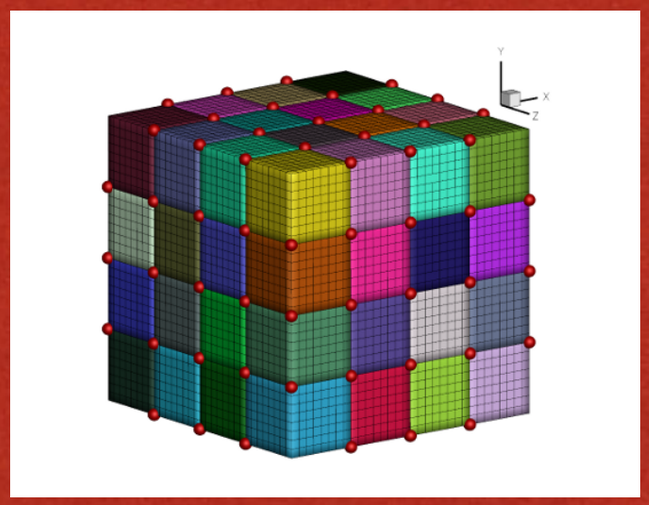


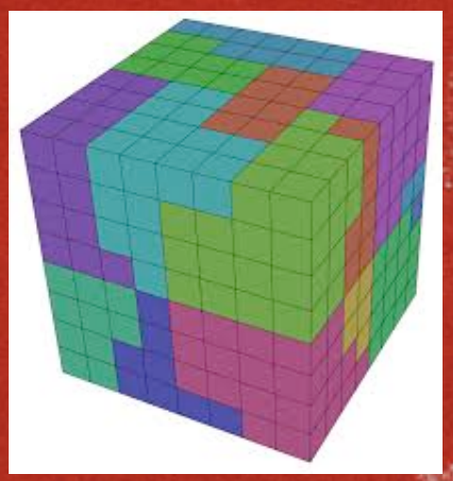
SAMPLE PROBLEM: DOMAIN DECOMPOSITION

- 1-d and 2-d: Lots of surface area
- 3-d: Generally more efficient for LARGE grids
- Irregular: Needs some careful programming!









SAMPLE PROBLEM: SPH -- SMOOTHED PARTICLE HYDRO

- We break a star into N=many "particles", really blobs of fluid that can overlap
- Each one has many neighbors, say M
- Each one has a position, velocity, etc.
- N operations are fast, but MN are not...
  - Do N position, velocity updates simultaneously on each processor, MN hydrodynamic force calculations in parallel.

XSEDE

- XSede:

- The Extreme Science and Engineering Discovery Environment (XSEDE) is the most advanced, powerful, and robust collection of integrated advanced digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data, and expertise.

- XSede is a set of supercomputers and associated resources that academics and others can use if they write a proposal for an allocation

- The largest machine is Stampede, #12 on the top 500 known supercomputers in the world

XSEDE STARTUPS

- The first step for an XSede allocation is to write a "Startup" proposal.
- Limits vary per machine, from up to 5,000 GPU-hours on Stanford's XStream machine and 50,000 core-hours on UC San Diego's Comet supercomputer. For reference, a year is 8,760 hours (+24 on leap years).
- Allocations are free (as in beer, not speech)!
- Startups require a 1-page proposal, with a primary investigator who is faculty or a postdoc at a university (corporations can also be eligible).

XSEDE ALLOCATIONS

- Startup allocations last a year, and are not usually renewable.
- After a year, the expectation is that if you want to continue, you will submit a full proposal:
  - 15 pages
  - Benchmarking results expected
  - Competitive, refereed process that requires science justification

FOR MORE INFO

- https://portal.xsede.org/allocations/startup
- https://portal.xsede.org/resource-monitor
- Homework -- think of a project related to your own that might require or benefit from supercomputing, and write up a sample paragraph
  - 1-2 sentences on the science problem you wish to study, with a reference or two
  - 1-2 sentences on methods, techniques, and parameter choices
  - 1 sentence estimating how much computer time it will require