Parallelization for Dummies

Graham Ganssle, Ph.D.

gra.m-gan.sl

High level computational langs

- C - lowest level here; memory leaks, heartburn, etc.
- python - huge number of ancillary functions through libs
- julia - fast like C, easy (scripting) like Python

“Automatic” parallelization

- C - OpenMP / CUDA

Mature, robust platforms.
Extreme efficiency, but you’ve got to be, like, an expert, man.

- Python - joblib / DisPy

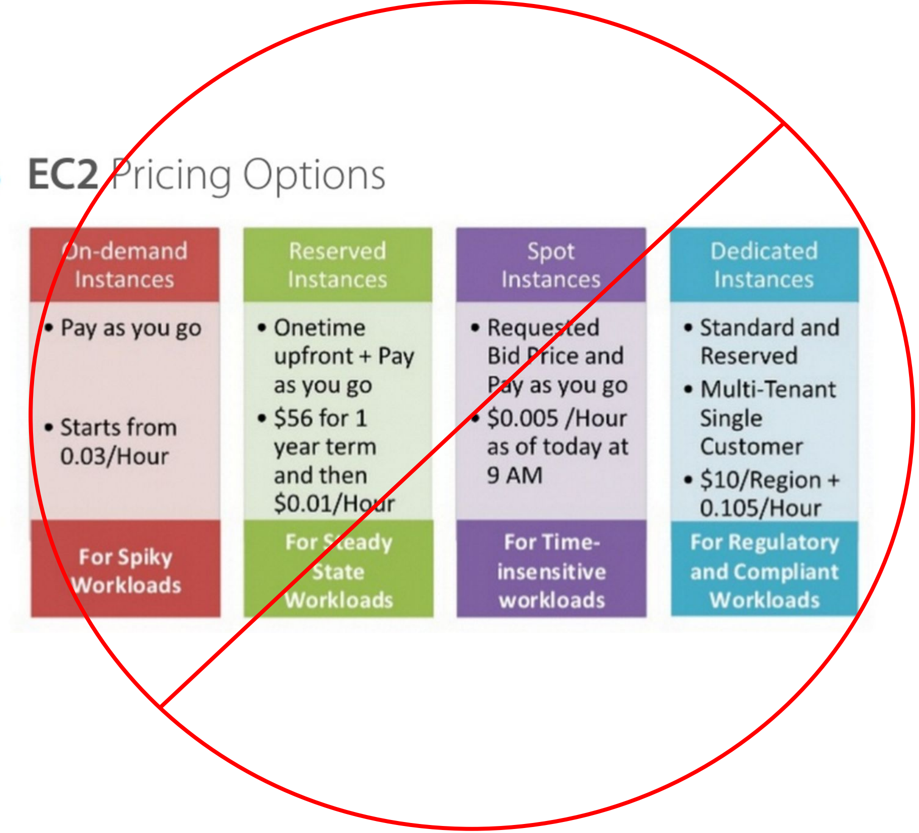
Plus a zillion other libs.
Stop working so hard. What you’re looking for exists already (it’s Python, duh).

- Julia - pmap() / @parallel

Super fun, but CPU parallel.
GPU solutions do exist; for example see the Julia CUDA package, CUDArt.jl.

AWS cloud rentals

- Cheap compute - EC2
- Repeatable builds - Vagrant
- Easy peasy storage - S3 / EBS



EC2 Pricing Options

On-demand Instances	Reserved Instances	Spot Instances	Dedicated Instances
• Pay as you go • Starts from 0.03/Hour	• Onetime upfront + Pay as you go • $56 for 1 year term and then $0.01/Hour	• Requested Bid Price and Pay as you go • $0.005 /Hour as of today at 9 AM	• Standard and Reserved • Multi-Tenant Single Customer • $10/Region + 0.105/Hour
For Spiky Workloads	For Steady State Workloads	For Time-insensitive workloads	For Regulatory and Compliant Workloads

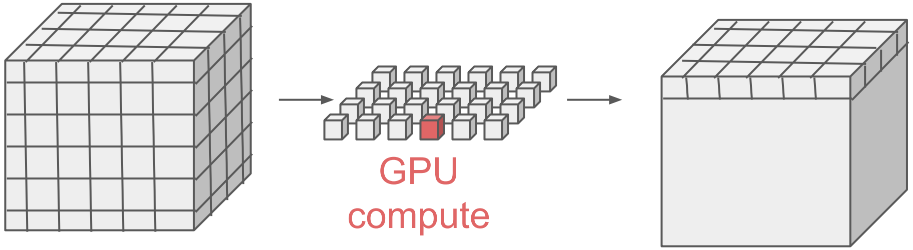
Requisite skills:

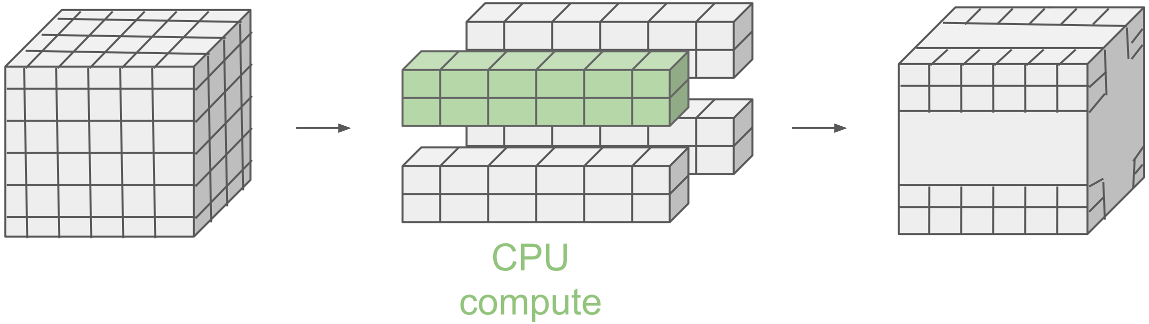
- Basic unix stuff
- SSH

S3 ~= 3 ¢/GB

CPU vs GPU parallelization

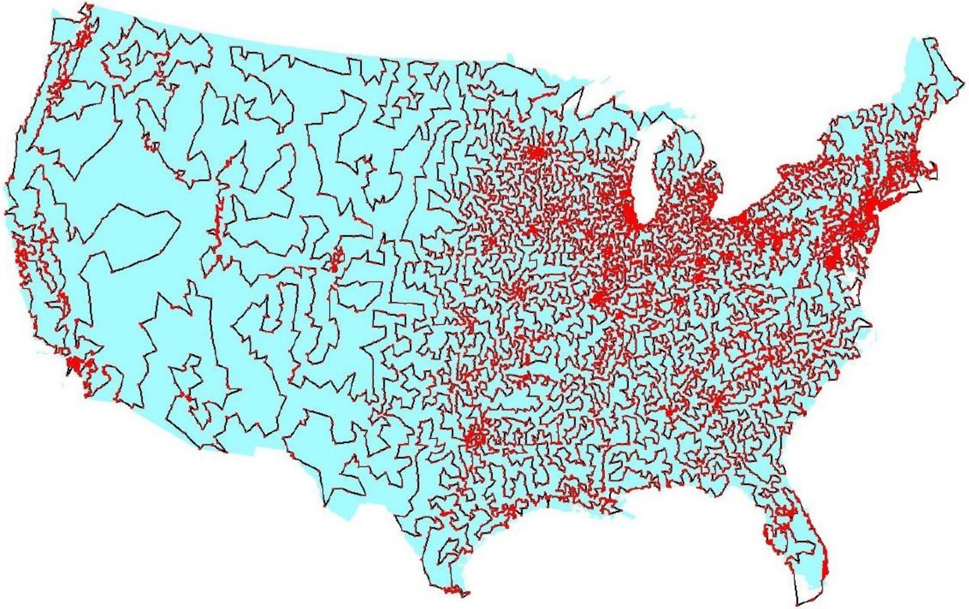
- GPU is awsm / highly efficient
    - Have to rebuild your algorithmic flow
- CPU is less efficient, but easy
    - Operates on basically the same algo





Application areas

- Inverse problems
- Optimization solutions
- Combinatorial algos

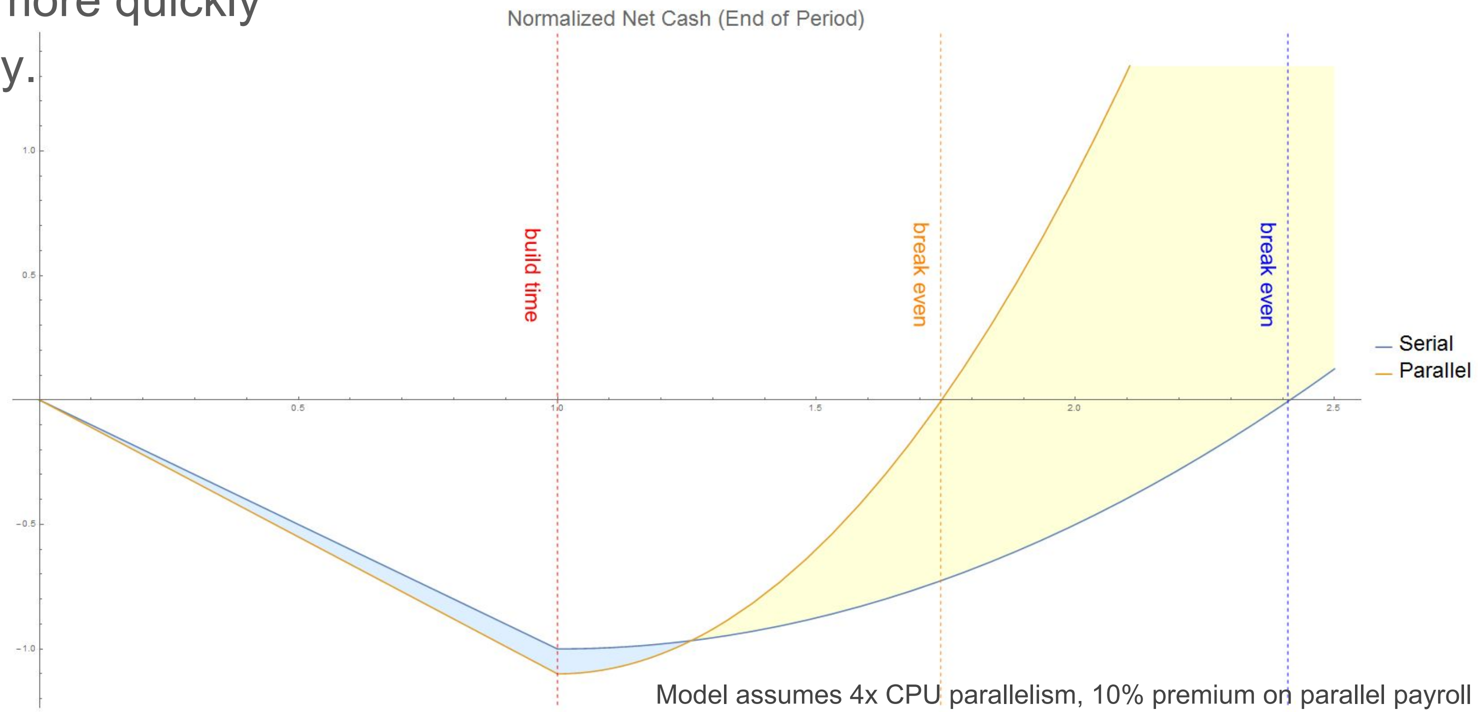


$$\frac{\partial J(\xi, \xi')}{\partial \xi} = \sum_{\gamma,\mu} S_{\mu}^{\dagger} S_{\mu}[u_{\gamma}(\xi) - u_{\gamma}^{'}(\xi^{'})]$$

Parallel model pays out **28%** faster than serial!

Profit!

1. Spend extra $ up front to CPU parallelize
2. Run things more quickly
3. Make money.



Model assumes 4x CPU parallelism, 10% premium on parallel payroll

Acknowledgements

Jaqcui Nelson

Sandstone Oil & Gas

R&D DESIGN

Quickie Trial Run

- ssh -i passwd/path os@ec2-xxx.yyy.zz.qq
- s3fs {bucket} folder/ -passwd_file as/df
- Julia
    - SharedArray()
    - @sync @parallel for (i = 1; i < x; i++)
- bash time {script}