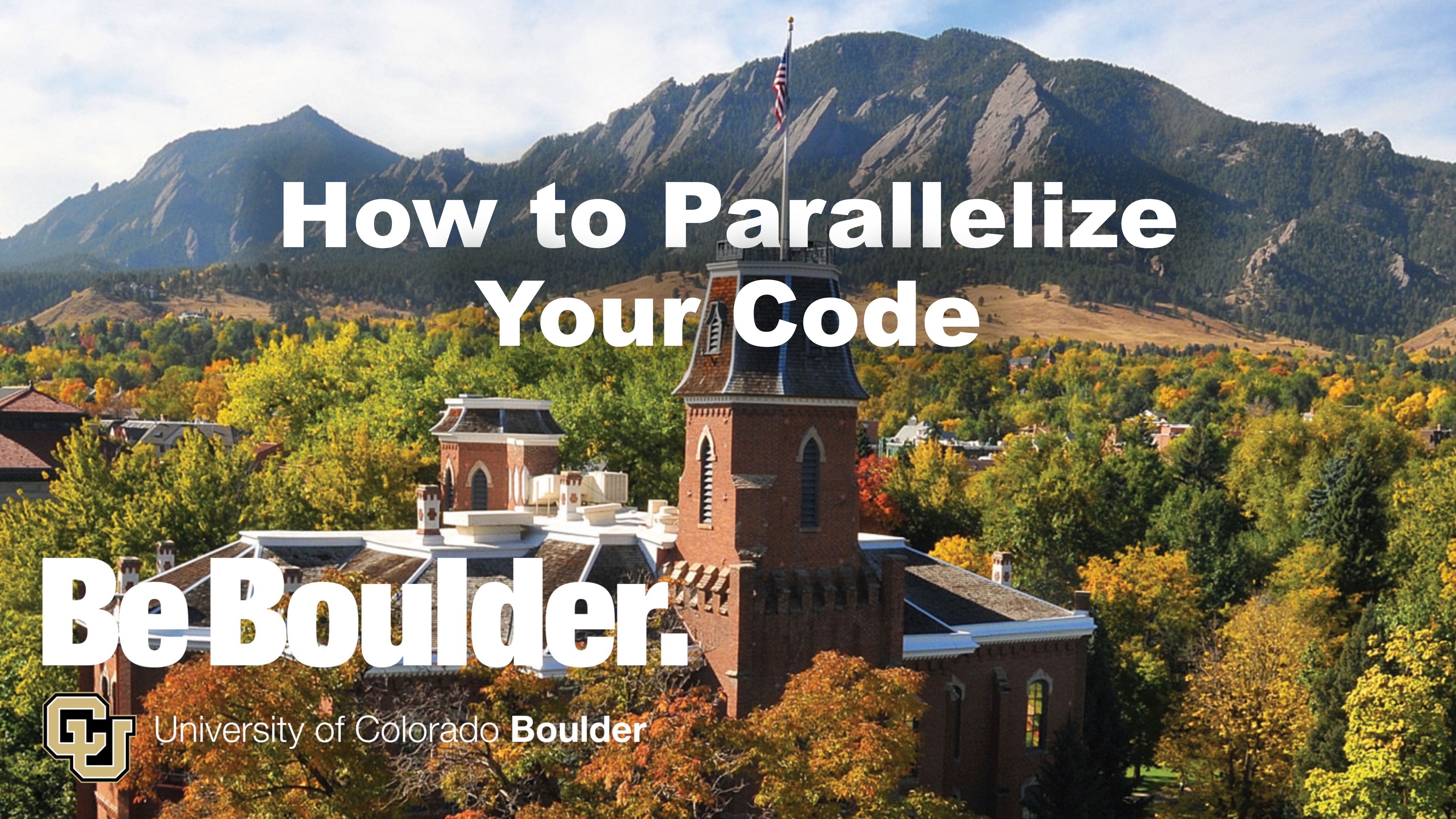
How to Parallelize Your Code



Be Boulder.

University of Colorado **Boulder**

Programming

- This is not a programming class!
- However, we’ve discussed a lot about both serial and parallel programs
- But haven’t spent a lot of time talking about *how* you can take serial program and parallelize it
- There are a lot of semantics depending on the code you write, but we can offer a few tips

General Parallelism

- You can parallelize code across cores on one node using shared memory or across nodes using distributed memory
- Generally it’s easier to parallelize using shared memory than distributed memory
- Parallel programming is either supported directly by the compiler or by libraries that will help you parallelize your code
- The idea here is to get your code to run simultaneously on multiple processors working on a subset of the problem which will save time

Why Parallelize?

- I don’t have time to learn how to parallelize my code!!
  - Usually saves time in the end!
- A program is a series of executed instructions
  - Your code is only going to run as fast as the slowest piece of code
  - If you’re running a relay race, you’re only as fast as the slowest person on your team
- A serial application does not take advantage of all the cores that are available on all current processors
  - You will never run faster

What can be parallelized?

- Code that can be executed independently of other instructions
  - In a loop, make sure iterations do not depend on other iterations

Dependency – not parallel:

```
for i=1:10
  a(i)=b(i)+c(i-1)
end
```

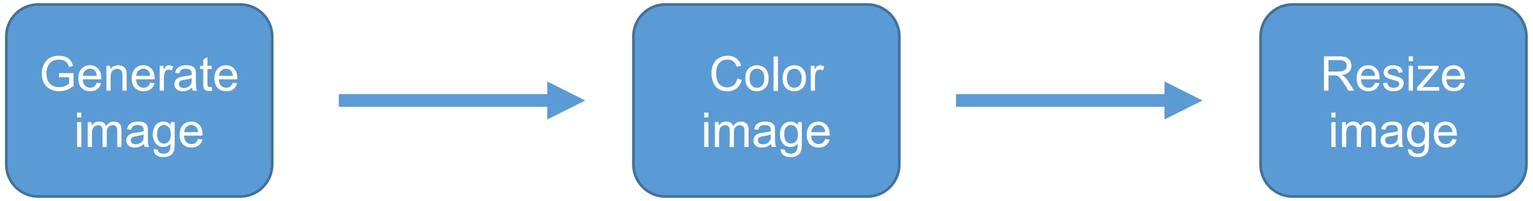
Dependency – not parallel:

```
for i=1:10
  a(i)=b(i)+c(i)
  d(i)=e(i)*a(i)
end
```

- Can you use different algorithms or functions instead that are intended or suitable for parallelization?

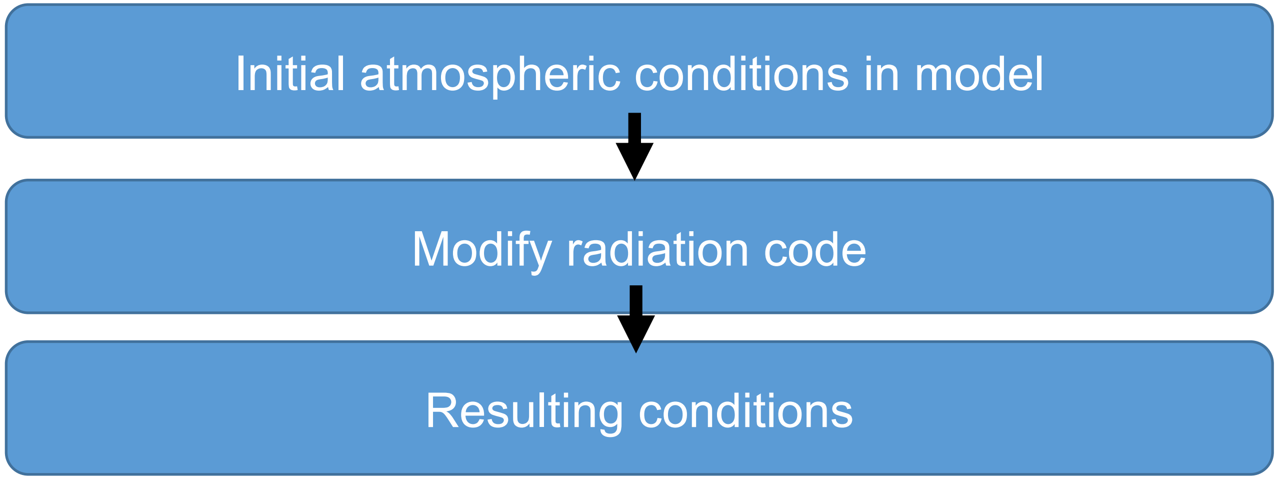
What code can be parallelized?

- Do you only need to pass the results one way?
  - Pipelining



Generate image → Color image → Resize image

- Each operation is performed on all of the data



Initial atmospheric conditions in model

Modify radiation code

Resulting conditions

How?

- Compiled languages (C++, Fortran, etc) use:
  - OpenMP: shared memory parallelization (within a node)
  - MPI: parallelize across nodes
    - Message passing and communication
  - Need to specify certain flags when compiling code
- Scripting languages (Python, R)
  - Packages that you can use that have functions that will allow you to parallelize based on certain bits of code
    - Functions to parallelize loops, for example
    - Functional programming, e.g. parallel apply