Replication & misc

BEACON CLASS
SEP 24TH

Today's topics

- Replication and reproducibility
- How much memory does Avida use?
- Saving highly evolved critters.

Replication vs Reproducibility

- Replication: identical results.
  - Same parameters => same results
- Reproducibility: similar results.
  - Similar parameters => similar results
  - (What's "similar"?)
- Corroboration:
  - Similar results seen in a different system

First: why did everyone get different results?

Avida is *stochastic*

- You really don't want to run exactly the same simulation every time you fix parameters…!
- But computers aren't (shouldn't be) random! So how??
- "Pseudo-random number generator"
  - Basically, take a number. Generate another number from it using a deterministic process, but make the process return a very *different* number from the first one, using complex math functions.

Avida is *stochastic*

- But this just pushes the problem back – how do you choose that first number differently for each run!?
- …take the time of day and use that as your random "seed", from which all your other numbers will be generated.
- This "seed" can then be used to replicate the run exactly.

Avida is stochastic

- So, if two of you had run Avida at the same microsecond, you should have gotten the same results!

Suppose…

- You cannot replicate the results in the 2003 paper.
- What are your options?

(Posit no evildoing)

Suppose…

- You find a copy of the source from 2003, and run the same program with the same parameters. And you get different results!
- Why might this happen?

Suppose…

- You get a copy of the source from 2003, and you run it hardware from 2003, and *still* get different results!?

What should our standards be?

- For *scientists*, reproducibility is extremely important.
- Replication … less so. It's very challenging to *exactly* replicate a given experimental situation.
- But, there is a pragmatic reason to think about replication, too.

Replication in computational science

- We have spent mucho time making sure that computers do the same thing *every time*, at the micro level.

- If you observe *unplanned* variation in a computational system, then:
  - You either are using one of the approximate subsystems, like floating point;
  - *Or* **you have a bug.**

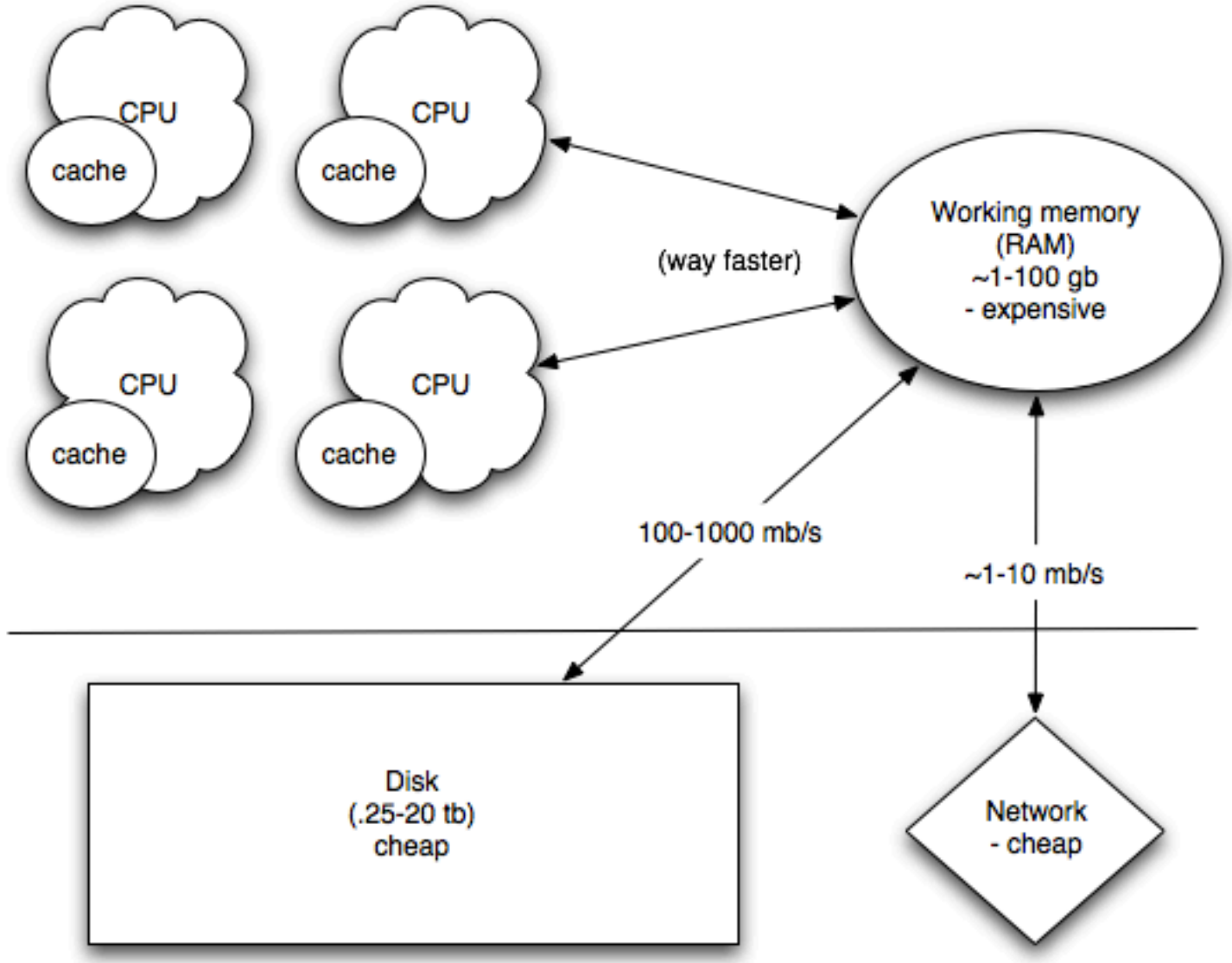
Replication in computational science

- So, *proximate to an experiment*, you should be able to exactly replicate a particular result from a computational experiment.
- Then, changing
  - Data sets
  - Hardware
  - Underlying software
- …may result in observed differences.

Engineering vs Science mindset

- Engineering mind set is aimed at *construction*. They care if it does as they intended it to do, and if they can reproduce the construction process.
- Scientists are trying figure out which characteristics are general and which ones aren't, by exploring the system.
- The two fields do connect quite a bit, especially in research situations; think complex systems.

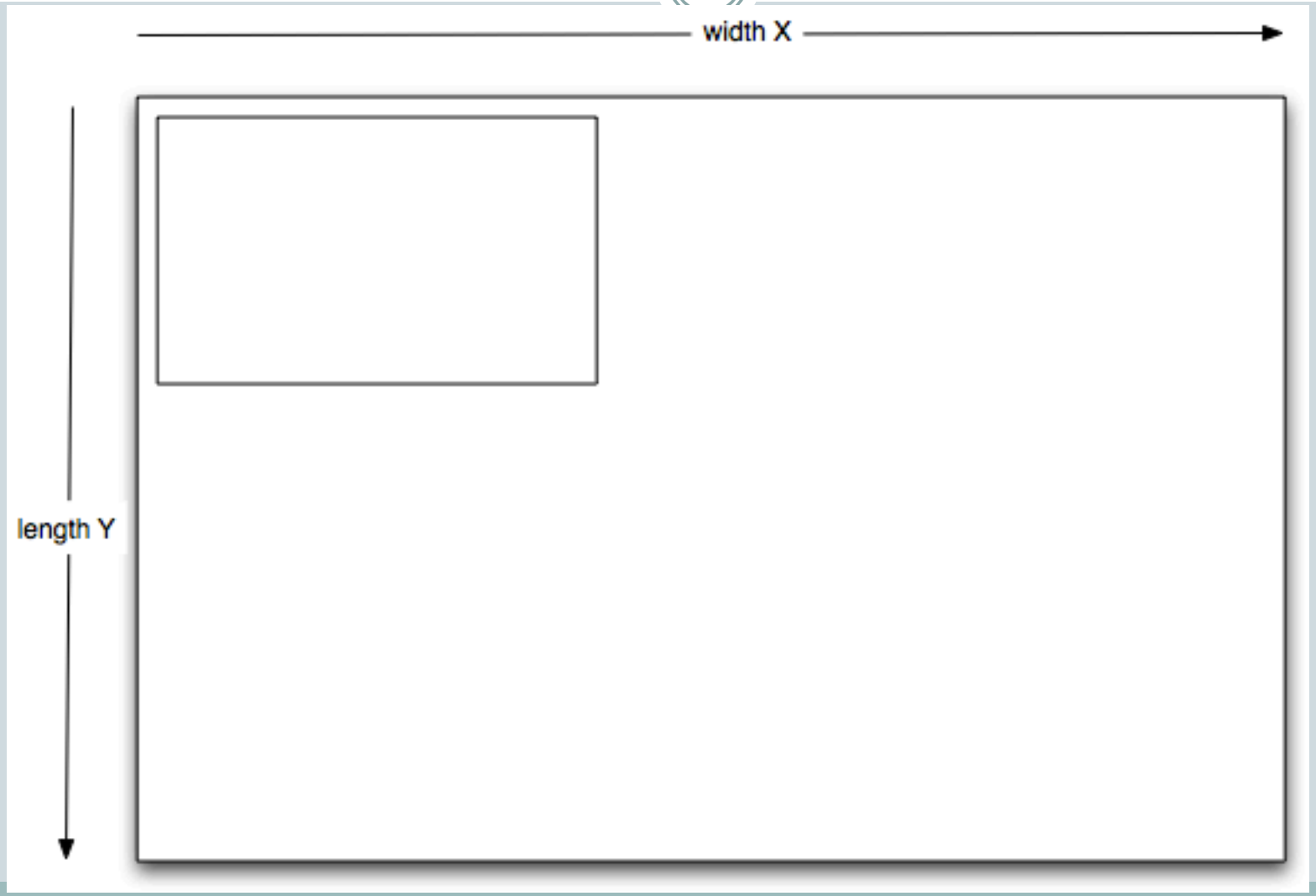
Computer architecture



How much memory do you think Avida uses?

- …and how does it scale with world size?

See? Lawn mowing == avida!



Doing your homework!

Avidian Gladitorial Combat