SAND2011-2161C

Scientific Simulations are Dead!...

Long Live Scientific Simulations!

March 30, 2011

Jay Lofstead

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

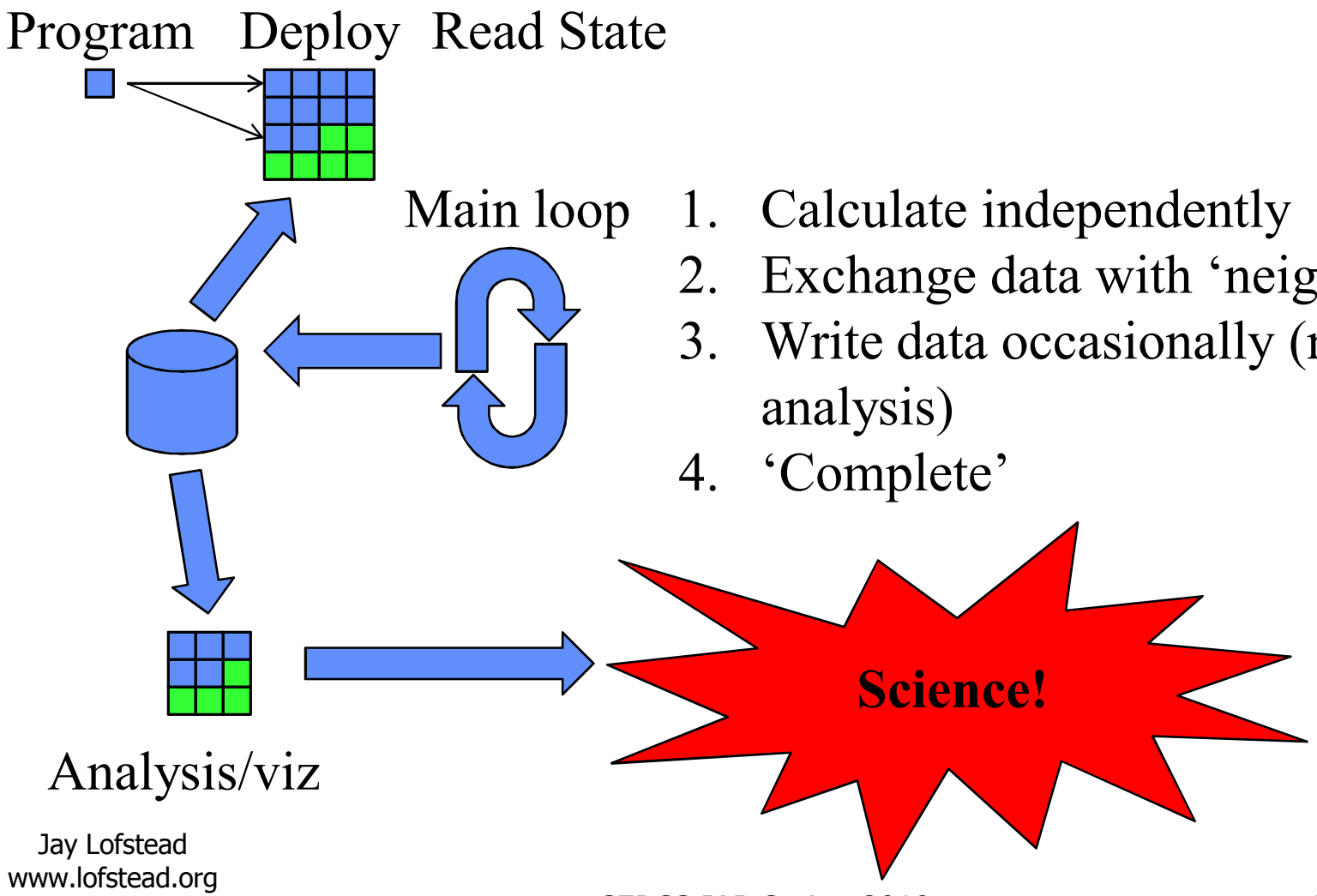
Why Are Scientific Simulations Dead?

- Tera-scale to peta-scale 'simple' evolution
  - Machines are generally lots more of the same
    - More cores, memory, storage, network, …
- Peta-scale to exa-scale is different
  - Co-processors expected, data movement must be avoided, storage can't hope to keep up, power consumption matters above all

Must meet performance with limited power use on very different architectures

Jay Lofstead
www.lofstead.org
gflofst@sandia.gov

Current Science Process (highlights)



1. Calculate independently
2. Exchange data with 'neighbors'
3. Write data occasionally (restarts, analysis)
4. 'Complete'

Jay Lofstead
www.lofstead.org
gflofst@sandia.gov

Where are the problems?

1. Moving data to storage
   - Not enough bandwidth, too energy intensive
2. Reading data from storage
   - Data too large, won't be available
3. Simulation writers assume can checkpoint enough to guard against failures preventing adequate progress
4. Provenance for publication support will be gone as we know it

Jay Lofstead
www.lofstead.org
gflofst@sandia.gov

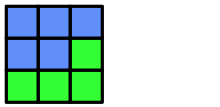
How Can We Fix This?

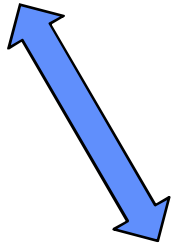
- **Current model: Simulation Driven Science**
  - **Design has focused on making simulations work well with managing data being a major driving force**
- **New model: Analysis or Discovery Driven Science**
  - **Make the data exploration and discovery process primary. All software and hardware must support the discovery process**

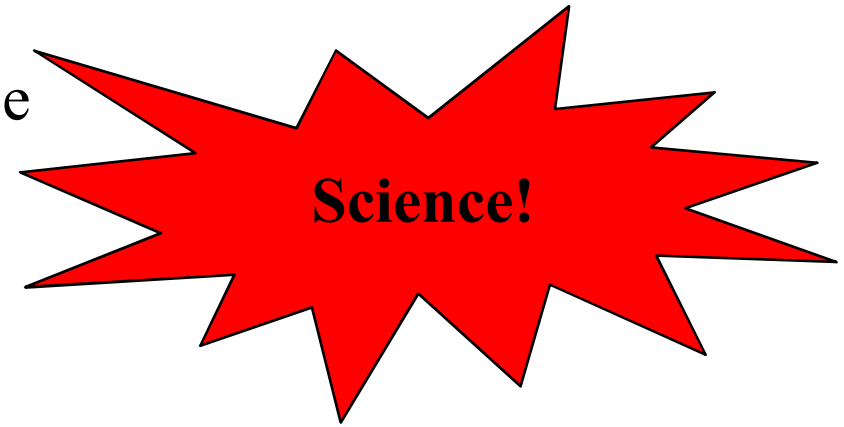
What Would This Look Like?

Analysis/viz

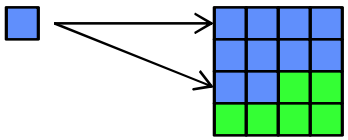
1. Configure parameters for exploration
2. Deploy code
3. Explore space

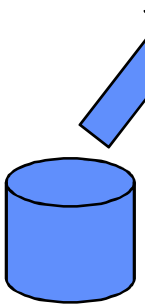




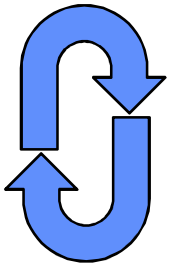


Program Deploy Read State





Main loop



1. Processes calculate independently in response to analysis requests
2. If data matches analysis request, send distilled data with provenance

Jay Lofstead
www.lofstead.org
gflofst@sandia.gov

We Have LOTS of Work to Do!

- Redesign from the top down
  - New top: visualization and analysis system
  - Everything must support discovery tasks
    - LOTS of simulation changes (math, cs, and physics)
    - How to support publications (or users) without data
    - Resilience techniques to work around failures
    - Intermediate data storage/processing to only output requested analysis summaries
    - Interactive workspaces
    - Automated discovery of data features (both backwards and forward)
    - Code coupling

Jay Lofstead
www.lofstead.org
gflofst@sandia.gov

We Are Working on All of It at Sandia

- Sandia proved scalability of massively parallel processing back in 1987
- Extensive work that led to the Cray XT3/4 architecture and software in use on the biggest machines in the world (Jaguar and Kraken, among others)
  - Developed original operating system (Catamount)
  - Developed portals networking protocol
- Still pushing boundaries of computing

Best Careers in the Next 10 Years

- Researcher – invent new stuff
- Internet/Wireless/Social system inventor – Capitalize on the integration of computing technology into all aspects of life
- Small scale, ultra high quality products manufacturing – food/restaurants, handmade goods
- Consulting – get paid to tell people what you know and understand as an outsider