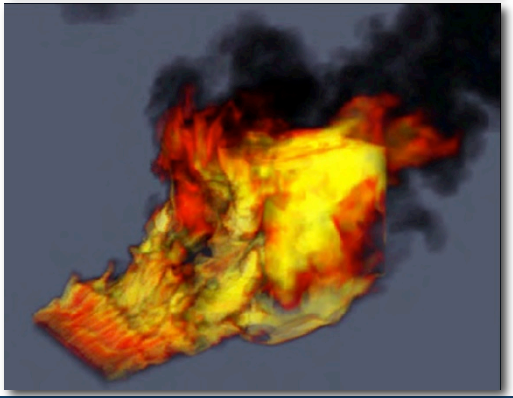


It's About SW Now, Not HW
It's About SW Now, Not Later

Michael A. Heroux
Senior Scientist
Sandia National Laboratories

Sandia National Laboratories

Exceptional service in the national interest

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

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.

Post-Exascale Computing:

- IS NOT primarily about HW.
- IS NOT primarily about Algorithms.
- IS primarily about software:
  - $\Delta A, \Delta S >> \Delta H$
  - Algs: Adequate R&D focus.
  - SW: Opportunities for improvement.

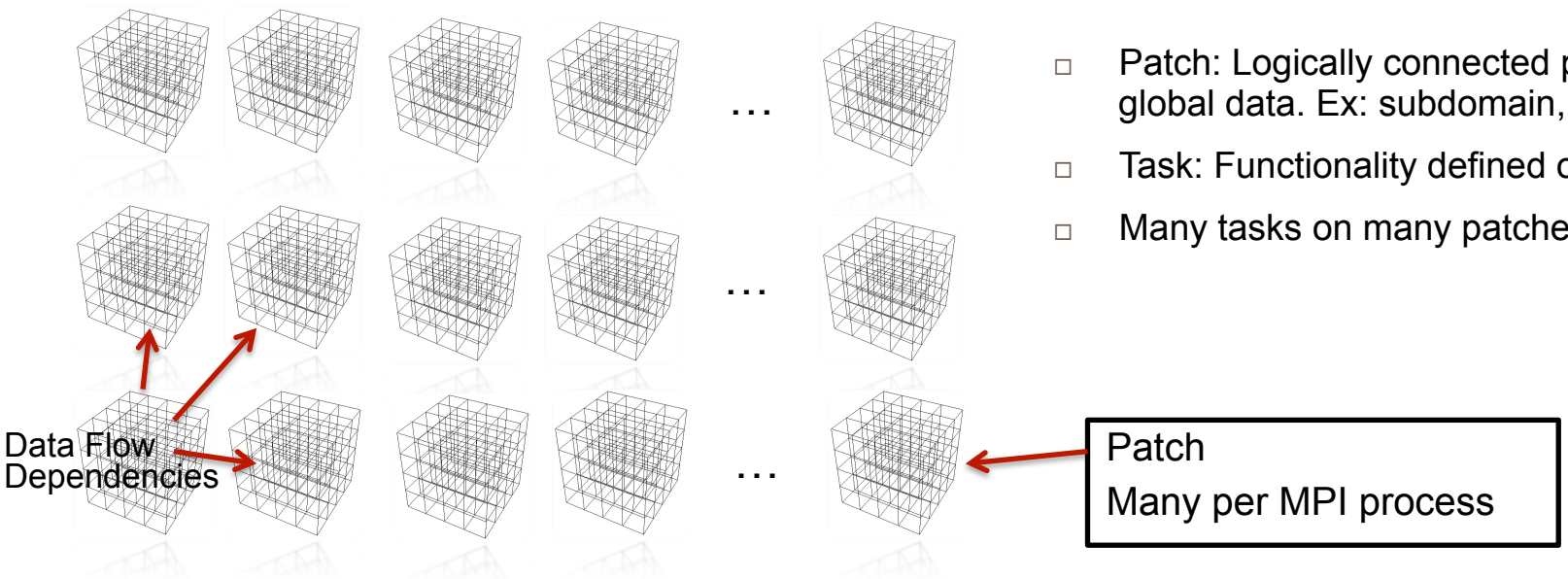
Quiz (Preview)

1. Which (a or b) is a proper first line of a Git commit message?
   a. Added test to computeConditionEstimate() that fixed Chris's problem.
   b. Add underflow test to computeConditionEstimate()
2. Which (a or b) is the best Agile workflow for Research SW?
   a. Scrum.
   b. Kanban.
3. Which entity can provide natural incentives for better SW?
   a. Publishers.
   b. Funding agencies.
   c. Both.
4. Which statement is true?
   1. Focus on reproducibility helps SW productivity and sustainability.
   2. Focus on SW productivity and sustainability helps reproducibility.
   3. Both.

Ubiquitous App Re-architecting

- Every Source Line of Code:
  - Refactored, re-encapsulated.
  - Can be done well, or poorly.
- New Algorithms:
  - Massive concurrency.
    - Task/Vector.
  - Deep memory hierarchies.
    - Nested parallelism.

Task-centric/Dataflow Application Architecture



- Patch: Logically connected portion of global data. Ex: subdomain, subgraph.
- Task: Functionality defined on a patch.
- Many tasks on many patches.

- Strengths:
  - Portable to many specific system architectures.
  - Separation of parallel model from implementation.
  - Domain scientists write sequential code within a parallel framework.
  - Supports traditional languages (Fortran, C).
  - Similar to SPMD in many ways.

- More strengths:
  - Well suited to emerging manycore systems.
  - Can exploit functional on-chip parallelism.
  - Can tolerate dynamic latencies.
  - Can support task/compute heterogeneity.

5

Multi-Team Science

- 10,000 X Petascale:
  - Enough fidelity at some point.
  - Multi-scale, multi-physics.
- Teams of teams.
  - Leveraging code.
  - Living with each other's good and bad practices.
  - Communicating.

Creating Incentives: Carrots & Sticks

- Opportunity:
  - Publishers & Funding Agencies can provide increased incentive for better behavior.
- Challenge:
  - Cultural change.
  - Increased cost: Fewer results, more cost (but better quality).
- Goal:
  - Increase incentives to improve reproducibility.
  - Improve SW quality, making reproducibility easier.
- Example:
  - TOMS Replicated Computational Results effort.
  - DOE SW Productivity and Sustainability Plan.
- Bottom Line:
  - You get published *if your results are reproducible.*
  - You get money *if you have a solid software plan.*

ACM TOMS Reproducible Computational Results (RCR) Process

- Submission: Optional (for now) RCR option.
- Standard reviewer assignment: Nothing changes.
- RCR reviewer assignment:
  - Concurrent with the first round of standard reviews
  - Known to and works with the authors during the RCR process.
- RCR process:
  - Multi-faceted approach.
- Publication:
  - Replicated Computational Results Designation.
  - The RCR referee acknowledged.
  - Review report appears with published manuscript.

TOMS RCR Process

Not reproducible, not published (eventually).

- Independent replication:
  - Transfer of or pointer to software given to RCR reviewer.
  - Guest account, access to software on author's system.
  - Detailed observation of the authors replicating the results.
- Review of computational results artifacts:
  - Results may be from a system that is no longer available.
  - Leadership class computing system.
  - In this situation:
    - Careful documentation of the process.
    - Software should have its own substantial verification process.

- First RCR paper in TOMS issue 41:3
  - Editorial introduction.
  - van Zee & van de Geijn, BLIS paper.
  - Referee report.
- Second: TOMS 42:1
  - Hogg & Scott next.
- Third: TOMS 42:4.
- Fourth ID'ed.

Cover letter excerpt from RCR candidate paper

Thank you for taking the time to consider our paper for your journal.

XXX has agreed to undergo the RCR process should the paper proceed far enough in the review process to qualify. ***To make this easier we have preserved the exact copy of the code used for the results (including additional code for generating detailed statistics that is not in the library version of the code).***

DOE SW Productivity and Sustainability Plan (SW PSP).

No viable SW plan, no money (eventually)

- Key Entities:
  - DOE Biological and Environmental Research (BER).
  - DOE Advanced Scientific Computing Research (ASCR)
  - IDEAS Scientific SW Productivity Project
- Milestone:
  - First-of-a-kind SW Productivity and Sustainability Plan.

DOE BER SW PSP Requirements (I)

- Describe overall SW development process.
  - Software lifecycle, testing, documentation and training.
- Development tools and processes:
  - source management, issue tracking, regression testing, SW distribution.
- Training and transition:
  - New and departing team members.
- Continuous process improvement:
  - Getting better at productivity and sustainability.

Interoperable Design of Extreme-scale Application Software (IDEAS)

Motivation

Enable *increased scientific productivity,* realizing the potential of extreme- scale computing, through *a new interdisciplinary and agile approach to the scientific software ecosystem.*

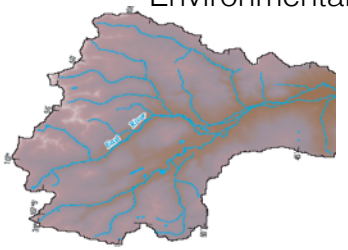
Objectives

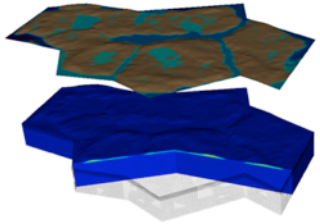
Address confluence of trends in hardware and increasing demands for predictive multiscale, multiphysics simulations.

Respond to trend of continuous refactoring with efficient agile software engineering methodologies and improved software design.

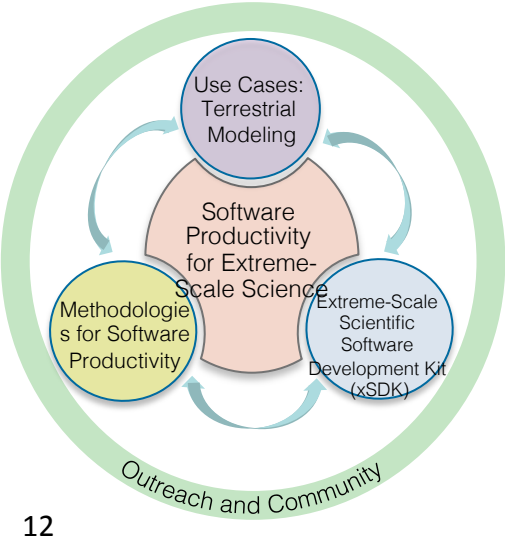
Impact on Applications & Programs

Terrestrial ecosystem *use cases tie IDEAS to modeling and simulation goals* in two Science Focus Area (SFA) programs and both Next Generation Ecosystem Experiment (NGEE) programs in DOE Biologic and Environmental Research (BER).





Approach



ASCR/BER partnership ensures delivery of both crosscutting methodologies and metrics with impact on real application and programs.

Interdisciplinary multi-lab team (ANL, LANL, LBNL, LLNL, ORNL, PNNL, SNL)

ASCR Co-Leads: Mike Heroux (SNL) and Lois Curfman McInnes (ANL)

BER Lead: David Moulton (LANL)

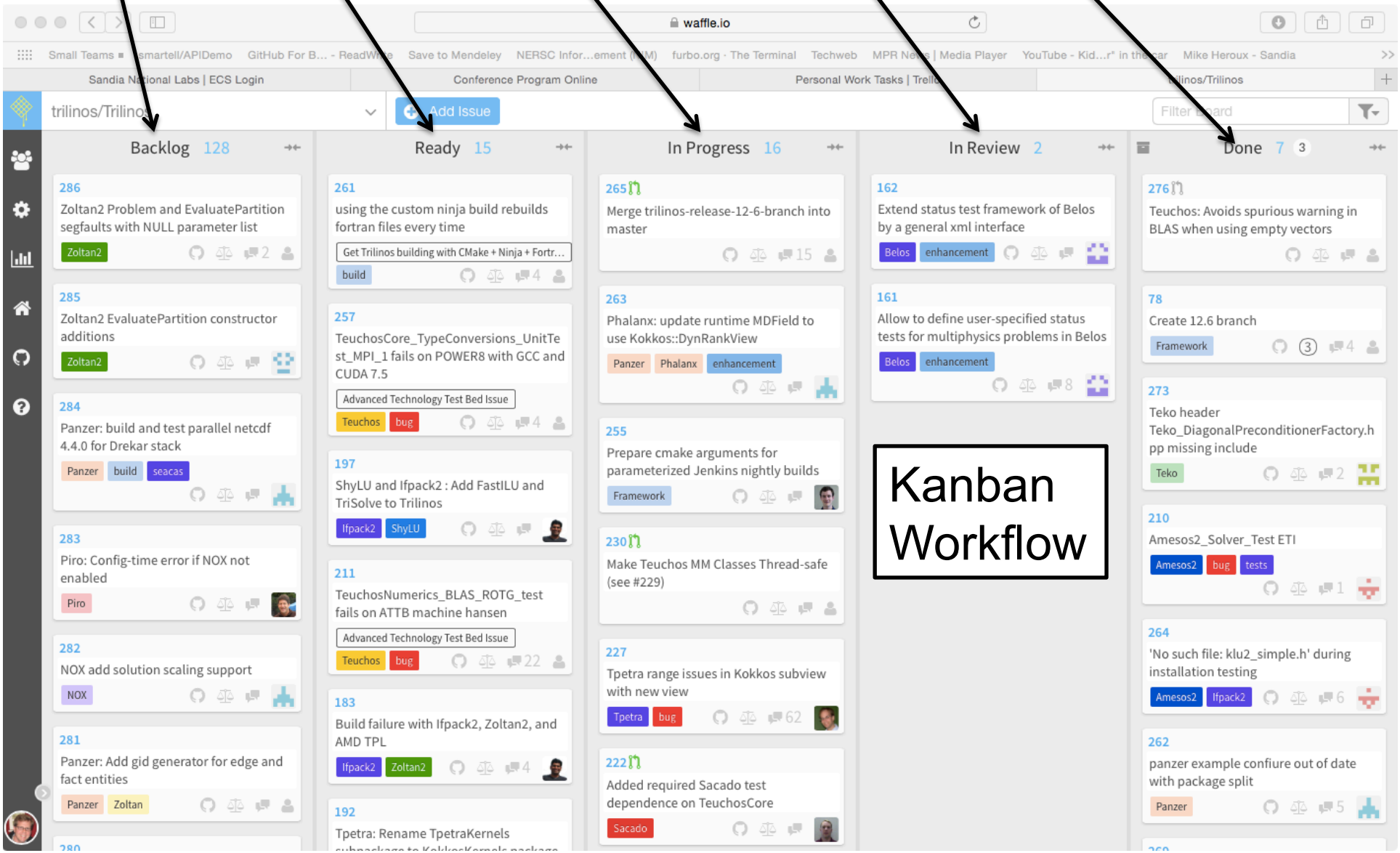
Topic Leads: David Bernholdt (ORNL) and Hans Johansen (LBNL)

Integration and synergistic advances in three communities deliver scientific productivity; outreach establishes a new holistic perspective for the broader scientific community.

www.ideas-productivity.org

Backlog | Ready | In Progress | In Review | Done

Sandia National Laboratories



Kanban Workflow

Message to This Audience

Be prepared to demonstrate reproducibility in order to publish.

Be prepared to demonstrate SW productivity, sustainability and improvement.

We have resources to help.

Final Thought: Commitment to Quality

Canadian engineers' oath (taken from Rudyard Kipling):

My Time I will not refuse;
my Thought I will not grudge;
my Care I will not deny
toward the honour, use,
stability and perfection of
any works to which I may be
called to set my hand.

http://commons.bcit.ca/update/2010/11/bcit-engineering-graduates-earn-their-iron-rings

Productivity++ Initiative

Sandia National Laboratories

Productivity++

- ✓ Traceable
- ✓ In Progress
- ✓ Sustainable
- ✓ Improved

Version 1.2

https://github.com/trilinos/Trilinos/wiki/Productivity---Initiative

Quiz

1. Which (a or b) is a proper first line of a Git commit message?
   a. Added test to computeConditionEstimate() that fixed Chris's problem.
   b. Add underflow test to computeConditionEstimate()
2. Which (a or b) is the best Agile workflow for Research SW?
   a. Scrum
   b. Kanban
3. What entities can provide natural incentives for better SW?
   a. Publishers
   b. Funding agencies
   c. Both
4. Which statement is true?
   1. Focus on reproducibility helps SW productivity and sustainability
   2. Focus on SW productivity and sustainability helps reproducibility
   3. Both