IDEAS:

Advancing Software Productivity for Extreme-scale Science

Lois Curfman McInnes, Argonne National Laboratory

on behalf of all IDEAS project members

ESS Working Group Meeting April 24, 2017

Trends and challenges

- **Fundamental trends:**
  - Disruptive hardware changes
    - Requires thorough algorithm/code refactoring
  - Demands for coupling, broader functionality
    - Multiphysics, multiscale, data analytics
- **Challenges:**
  - Need refactoring: Really, continuous change
  - Modest funding for app development: No monolithic apps
  - Requirements are unfolding, evolving, not fully known *a priori*
- **Opportunities:**
  - Better design, software practices, and tools are available
  - Better software architectures: toolkits, libraries, frameworks
  - Open-source software, community collaboration

Interoperable Design of Extreme-scale Application Software (IDEAS)

IDEAS productivity

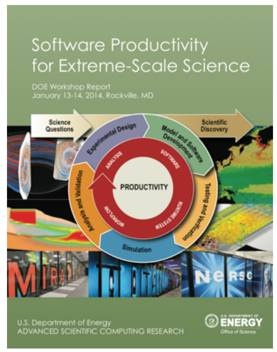
Motivation

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

Objectives

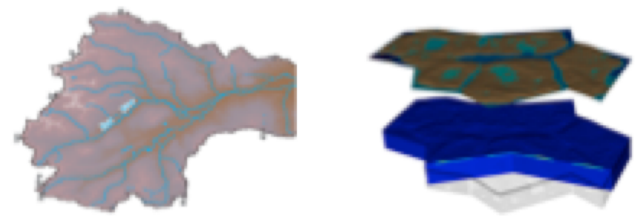
Address confluence of trends in hardware and increasing demands for predictive multiscale, multiphysics simulations.

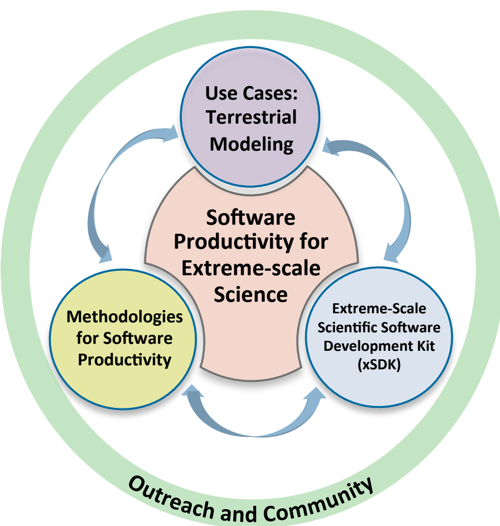
Respond to trend of continuous refactoring with efficient agile software engineering methodologies and improved software design.



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)

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

U.S. DEPARTMENT OF ENERGY
Office of Science

IDEAS project team

DOE Program Managers

ASCR: Thomas Ndousse-Fetter
BER: Paul Bayer, David Lesmes

IDEAS: Interoperable Design of Extreme-scale Application Software

ASCR Co-Leads: Mike Heroux (SNL) and Lois Curfman McInnes (ANL)
BER Lead: J. David Moulton (LANL)

Executive Advisory Board

John Cary (Tech-X)
Mike Glass (SNL)
Susan Hubbard (LBNL)
Doug Kothe (ORNL)
Sandy Landsberg (DOD)

BER Use Cases

Lead: J. David Moulton (LANL)

Carl Steefel (LBNL) *1
Scott Painter (ORNL) *2
Reed Maxwell (CSM) *3
Glenn Hammond (SNL)
Tim Scheibe (PNNL)
Laura Condon (CSM)
Ethan Coon (LANL)
Dipankar Dwivedi (LBNL)
Jeff Johnson (LBNL)
Eugene Kikinzon (LANL)
Sergi Molins (LBNL)
Steve Smith (LLNL)
Erica Woodburn (LBNL)
Carol Woodward (LLNL)
Xiaofan Yang (PNNL)

Methodologies for Software Productivity

Lead: Mike Heroux (SNL)

Roscoe Bartlett (SNL)
Mark Berrill (ORNL)
Anshu Dubey (ANL)
Todd Gamblin* (LLNL)
Osni Marques (LBNL)
Andrew Salinger* (SNL)
Jason Sarich (ANL)

Extreme-scale Scientific Software Development Kit (xSDK)

Lead: Lois Curfman McInnes (ANL)

Alicia Klinvex (SNL)
Sherry Li (LBNL)
Vijay Mahadevan (ANL)
Daniel Osei-Kuffuor (LLNL)
Barry Smith (ANL)
Keita Teranishi (SNL)
Mathew Thomas (PNNL)
Jim Willenbring (SNL)
Ulrike Yang (LLNL)

Outreach and Community

Lead: David Bernholdt (ORNL)

Lisa Childers* (ALCF)
Rebecca Hartman-Baker* (NERSC)
Judy Hill* (OLCF)
Bill Spotz* (SNL)

Crosscutting Lead: Hans Johansen (LBNL)

April 2017

* Liaison
*1 *2 *3: Leads: Use Cases 1, 2, 3

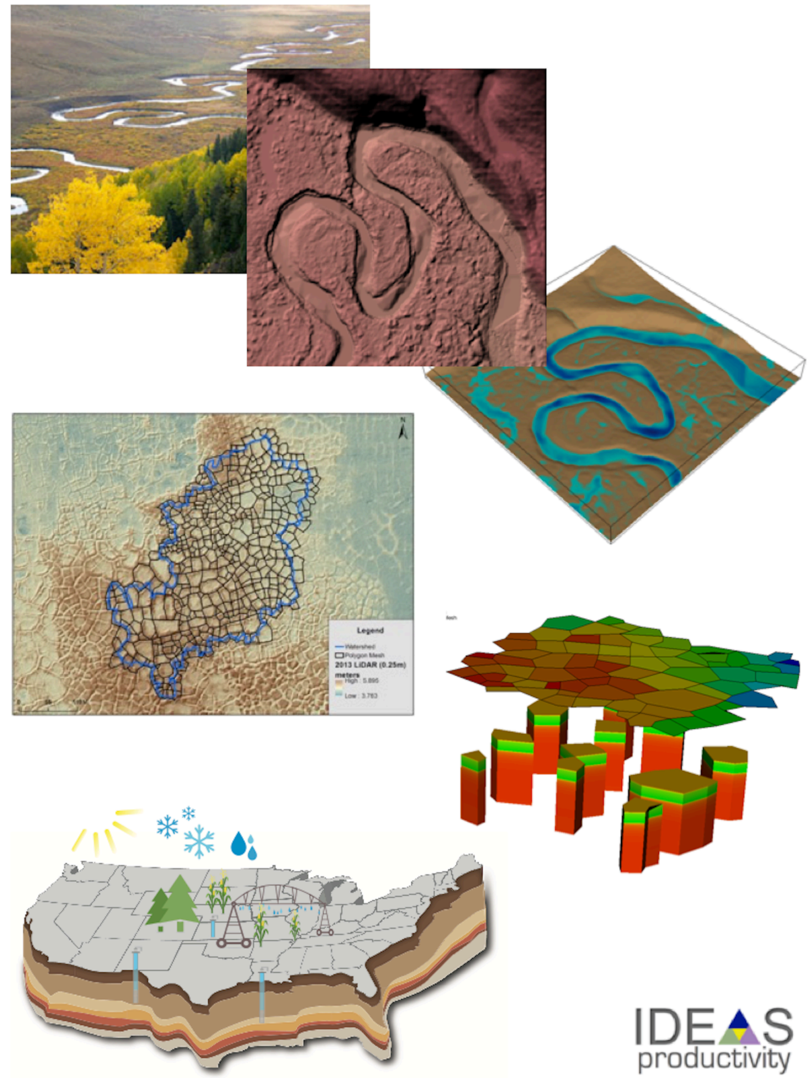
BER Terrestrial Programs: SFAs, NGEE, CLM, ACME

DOE Extreme-scale Programs: Exascale Co-Design, Exascale Roadmap, ASCR Math & CS, SciDAC

DOE Computing Facilities: ALCF, NERSC, OLCF

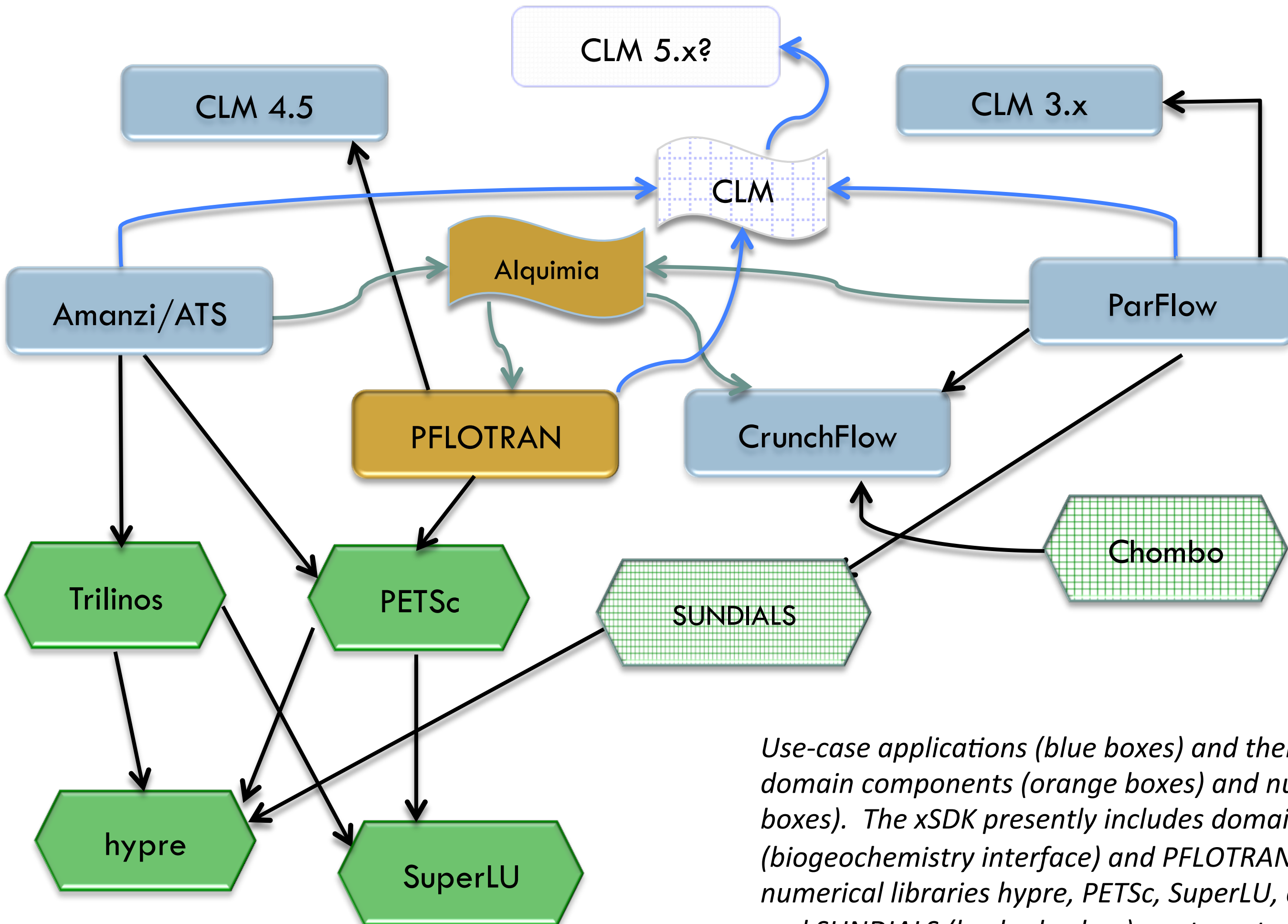
Use cases: Multiscale, multiphysics representation of watershed dynamics

- **Use Case 1:** Hydrological and biogeochemical cycling in the Colorado River System
- **Use Case 2:** Thermal hydrology and carbon cycling in tundra at the Barrow Observatory
- **Use Case 3:** Hydrologic, land surface, and atmospheric process coupling over CONUS
- **Leverage & complement SBR, TES programs:**
  - LBNL and PNNL SFAs
  - NGEE Arctic and Tropics
- **Approach:**
  - Leverage existing open source apps
  - Improve software development practices
  - Targeted refactoring of interfaces, data structures, and key components to facilitate interoperability
  - Modernize management of multiphysics integration and multiscale coupling

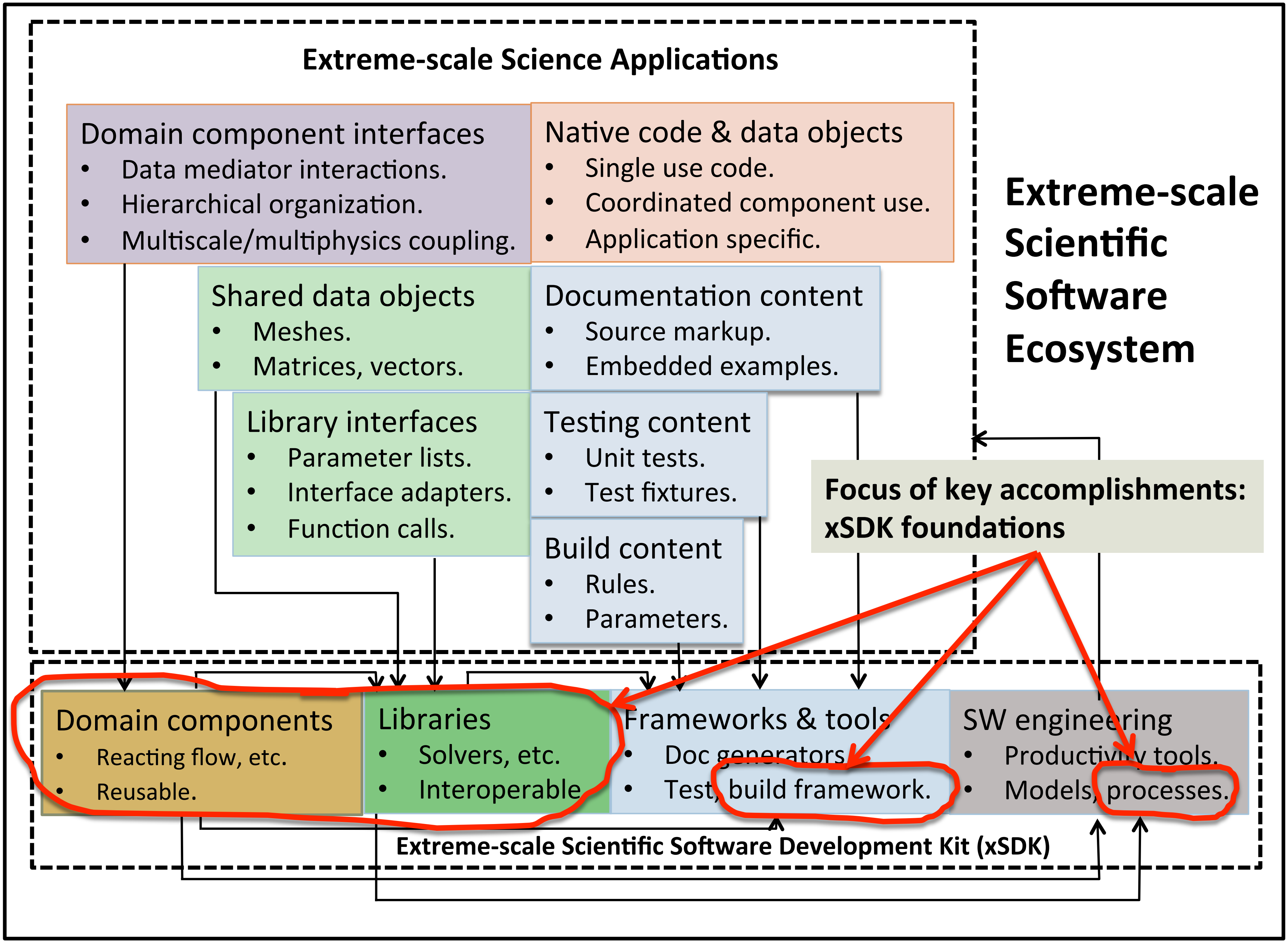


IDEAS productivity

Next-generation subsurface modeling requires the combined use of independent packages

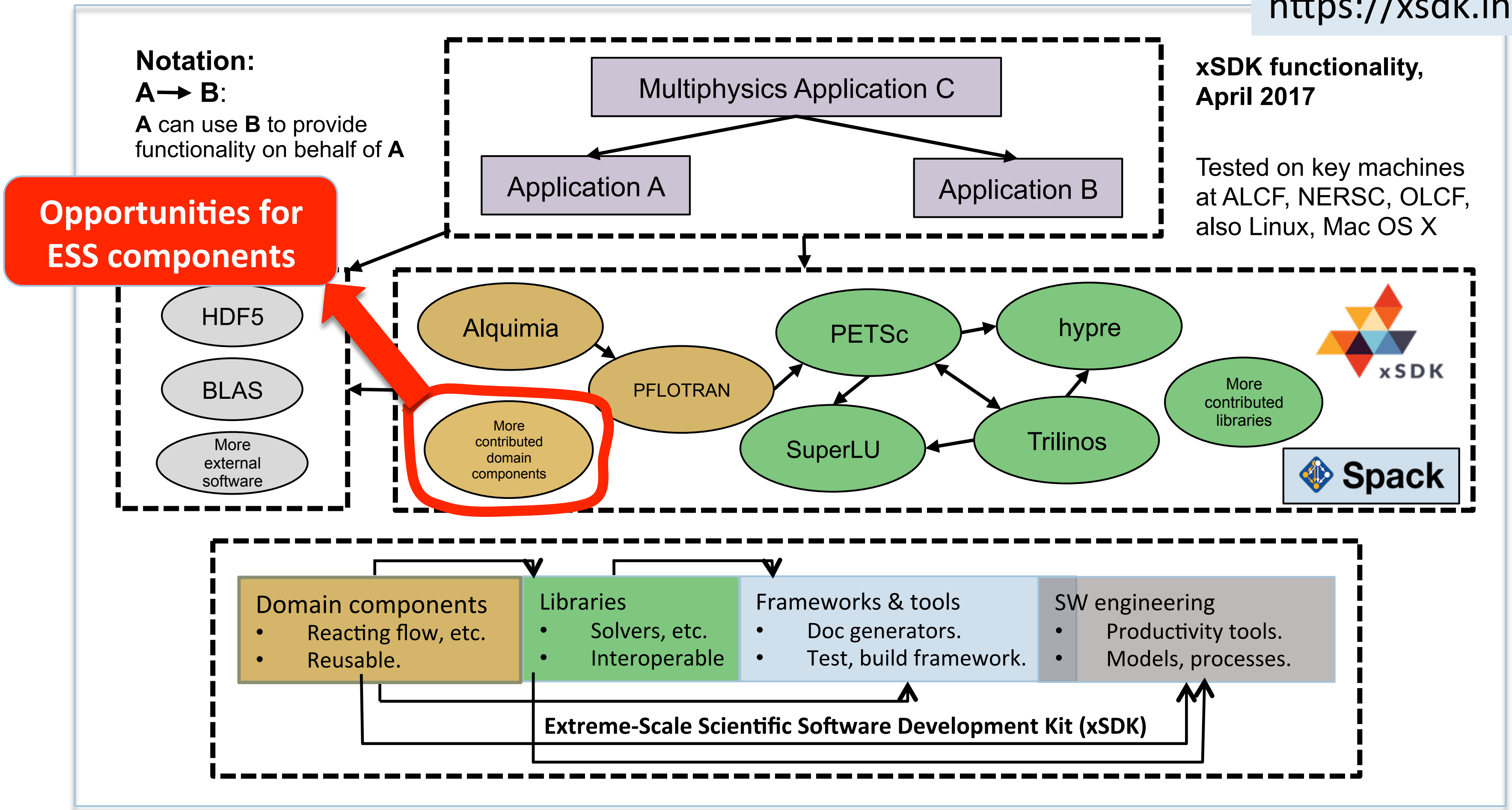


Use-case applications (blue boxes) and their present usage of SDK domain components (orange boxes) and numerical libraries (green boxes). The xSDK presently includes domain components Alquimia (biogeochemistry interface) and PFLOTRAN (subsurface flow) and the numerical libraries hypre, PETSc, SuperLU, and Trilinos. CLM, Chombo, and SUNDIALS (hashed colors) are targeted for later inclusion.


Extreme-scale Science Applications
Domain component interfaces
• Data mediator interactions.
• Hierarchical organization.
• Multiscale/multiphysics coupling.
Native code & data objects
• Single use code.
• Coordinated component use.
• Application specific.
Shared data objects
• Meshes.
• Matrices, vectors.
Documentation content
• Source markup.
• Embedded examples.
Library interfaces
• Parameter lists.
• Interface adapters.
• Function calls.
Testing content
• Unit tests.
• Test fixtures.
Build content
• Rules.
• Parameters.
Extreme-scale Scientific Software Ecosystem
Focus of key accomplishments: xSDK foundations
Domain components
• Reacting flow, etc.
• Reusable.
Libraries
• Solvers, etc.
• Interoperable
Frameworks & tools
• Doc generators.
• Test, build framework.
SW engineering
• Productivity tools.
• Models, processes.
Extreme-scale Scientific Software Development Kit (xSDK)

xSDK release 0.2.0: Packages can be readily used in combination by multiphysics, multiscale apps

https://xsdk.info



Ref: *xSDK Foundations: Toward an Extreme-scale Scientific Software Development Kit,* Bartlett et al, Feb 2017, https://arxiv.org/abs/1702.08425, to appear in **Supercomputing Frontiers and Innovations**.

IDEAS productivity

xSDK community policies

xSDK

Draft 0.3, Dec 2016

<u>xSDK compatible package</u>: Must satisfy mandatory xSDK policies:

M1. Support xSDK community GNU Autoconf or CMake options.

M2. Provide a comprehensive test suite.

M3. Employ user-provided MPI communicator.

M4. Give best effort at portability to key architectures.

M5. Provide a documented, reliable way to contact the development team.

M6. Respect system resources and settings made by other previously called packages.

M7. Come with an open source license.

M8. Provide a runtime API to return the current version number of the software.

M9. Use a limited and well-defined symbol, macro, library, and include file name space.

M10. Provide an accessible repository (not necessarily publicly available).

M11. Have no hardwired print or IO statements.

M12. Allow installing, building, and linking against an outside copy of external software.

M13. Install headers and libraries under <prefix>/include/ and <prefix>/lib/.

M14. Be buildable using 64 bit pointers. 32 bit is optional.

Also specify **recommended policies**, which currently are encouraged but not required:

R1. Have a public repository.

R2. Possible to run test suite under valgrind in order to test for memory corruption issues.

R3. Adopt and document consistent system for error conditions/exceptions.

R4. Free all system resources it has acquired as soon as they are no longer needed.

R5. Provide a mechanism to export ordered list of library dependencies.

<u>xSDK member package</u>: Must be an xSDK-compatible package, *and* it uses or can be used by another package in the xSDK, and the connecting interface is regularly tested for regressions.

We welcome feedback. What policies make sense for ESS apps and packages?

https://xsdk.info/policies

IDEAS productivity

Resources for software productivity & sustainability—key element of overall scientific productivity

Approach: Collaborate with the community to curate, create, & disseminate software methodologies, processes, and tools that lead to improved scientific software

Impact:

Better: Science, portability, robustness, composability
Faster: Execution, development, dissemination
Cheaper: Fewer staff hours and lines of code

CSE17 content available via https://ideas-productivity.org/events
Transitioning summer 2017 to:

better scientific software

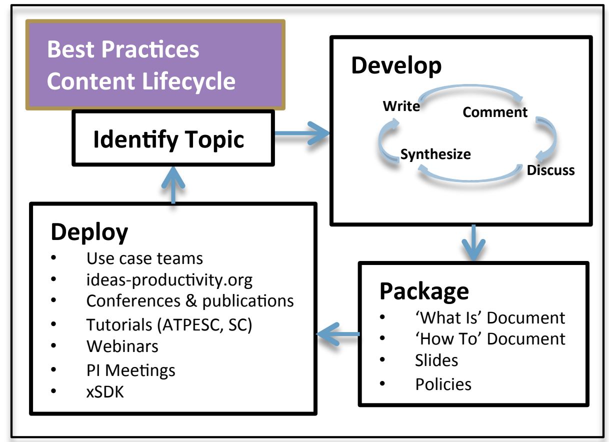
Coming soon: new web-based hub for collaborative content development & delivery

Modern learning theory:

Build from knowledge base: Elaboration and models

Vast body of SE content from broad community

Learn, adapt, adopt, assimilate



CSE17 Tutorial: *CSE Collaboration through Software: Improving Software Productivity & Sustainability*

- *Why Effective Software Practices Are Essential for CSE Projects*
- *An Introduction to Software Licensing*
- *Better (Small) Scientific Software Teams*
- *Improving Reproducibility through Better Software Practices*
- *Testing of HPC Scientific Software*

***What Is and HowTo* docs**: brief sketches of best practices

- *What Is CSE Software Productivity?*
- *What Is Software Configuration?*
- *How to Configure Software*
- *What Is Performance Portability?*
- *How to Enable Performance Portability*
- *What Are Software Testing Practices?*
- *How to Add and Improve Testing in a CSE Software Project*
- *What Is Good Documentation?*
- *How to Write Good Documentation*
- *What Are Interoperable Software Libraries?*
- *What Is Version Control?*
- *How to Do Version Control with Git*

[More topics under development]

Next steps

EXASCALE COMPUTING PROJECT

xSDK4ECP: Develop **community policies** and **interoperability layers** among numerical packages as needed by ECP scientific applications

xSDK

- Coordinated use of on-node resources
- Integrated execusion
  - Control inversion, adaptive execution strategies
- Coordinated and sustainable documentation, testing, packaging, and deployment

Current xSDK packages:

- **Numerical libraries: hypre, PETSc, SuperLU, Trilinos**
- **Domain components: Alquimia, PFLOTRAN**

Packages working toward xSDK compatibility:

- **Chombo, SUNDIALS, ALExa (**AMP, DTK, TASMANIAN)
- **Dense linear algebra packages:** MAGMA, PLASMA, DPLASMA, ScaLAPACK, LAPACK

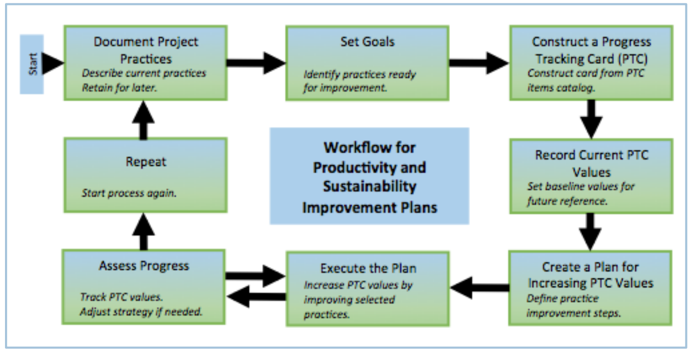
xSDK4ECP Team:

Michael Heroux (SNL), Co-Lead PI
Lois Curfman McInnes (ANL), Co-Lead PI
Jim Demmel (UC Berkeley), Co-PI
Jack Dongarra (UTK), Co-PI
Xiaoye Sherry Li (LBNL), Co-PI
Carol Woodward (LLNL), Co-PI
Ulrike Meier Yang (LLNL), Co-PI
Satish Balay (ANL)
David Gardner (LLNL)
Alicia Klinvex (SNL)
Piotr Luszczek (UTK)
Slaven Peles (LLNL)
Ben Recht (UC Berkeley)
Jacob Schroder (LLNL)
Barry Smith (ANL)
Keita Teranishi (SNL)
Jim Willenbring (SNL)

IDEAS-ECP: Collaborate with ECP app teams to understand productivity bottlenecks and improve practices

Productivity and Sustainability Improvement Planning Tools:

Helping a software team to increase software quality while decreasing the effort, time, and cost to develop, deploy, maintain, and extend software over its intended lifetime. https://github.com/betterscientificsoftware/PSIP-Tools



better scientific software

Summer 2017: New web-based hub for scientific software improvement exchange

IDEAS-ECP Team:

Michael Heroux (SNL), Co-Lead PI
Lois Curfman McInnes (ANL), Co-Lead PI
David Bernholdt (ORNL), Co-PI, Outreach Lead
Todd Gamblin (LLNL), Co-PI
Osni Marques (LBNL), Co-PI
David Moulton (LANL), Co-PI
Boyana Norris (Univ of Oregon), Co-PI
Satish Balay (ANL)
Roscoe Bartlett (SNL)
Anshu Dubey (ANL)
Rinku Gupta (ANL)
Christoph Junghans (LANL)
Alicia Klinvex (SNL)
Reed Milewicz (SNL)
Mark Miller (LLNL)
Elaine Raybourn (SNL)
Barry Smith (ANL)
Louis Vernon (LANL)
Greg Watson (ORNL)
James Willenbring (SNL)
Lisa Childers (ALCF)
Rebecca Hartman-Baker (NERSC)
Judy Hill (OLCF)
Hai Ah Nam (LANL)
Jean Shuler (LLNL)

Facilities liaisons: Lisa Childers (ALCF), Rebecca Hartman-Baker (NERSC), Judy Hill (OLCF), Hai Ah Nam (LANL), Jean Shuler (LLNL)

Opportunities for ESS collaboration

ESS software in xSDK

We are actively soliciting contributions to the xSDK and feedback on draft xSDK community policies. See https://xsdk.info/faq

- **xSDK compatible package**
  - Must satisfy mandatory xSDK policies
- **xSDK member package**
  - Must be an xSDK-compatible package, *and* it uses or can be used by another package in the xSDK, and the connecting interface is regularly tested for regressions

Why participate?

- Improved code quality, usability, access, sustainability
- Inform potential users that an xSDK member package can be easily used with others
- Foundation for work on performance portability, deeper levels of package interoperability

Opportunities for ESS components

We welcome feedback. What policies make sense for ESS apps and packages?

ESS contributions to better scientific software

- Collaborate to curate, create, & disseminate resources that lead to better ESS software
  - Key element of overall scientific productivity
  - ESS focus topic: **Team-of-teams**: collaborating across teams for multiscale, multiphysics ESS modeling and simulation
- Environmental System Science Community landing page tailored specifically to ESS priorities
  - ESS community determine emphasis & customization