OAK RIDGE National Laboratory

# COLABS: Collaboration for Better Software (for Science)

Anshu Dubey (ANL), David E. Bernholdt (ORNL), Dan Gunter (LBNL), Kevin Harms (ALCF), Bronson Messer (OLCF), John MacAuley (Esnet), and Richard Gerber (NERSC)

https://colabs-science.github.io

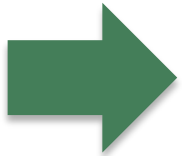
ORNL is managed by UT-Battelle LLC for the US Department of Energy

This work is supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research.

U.S. DEPARTMENT OF ENERGY

# The Fundamental Premise of COLABS

**Good scientific process**
**requires**
**good software practices**



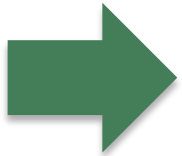
**Good software practices**
**increase**
**software sustainability**

**Good software practices**
**increase**
**scientific productivity**

**Software sustainability**
**increases**
**scientific productivity**

# The Fundamental Premise of COLABS

**Good scientific process requires good software practices**



**Good software practices increase software sustainability**

**Good software practices increase scientific productivity**

**Software sustainability increases scientific productivity**

***Science through computing is, at best, as good as the software that produces it***

# COLABS

**Comprehensive services in support of scientific software stewardship and sustainability**

# COLABS

**Comprehensive services in support of scientific software stewardship and sustainability**

*And…*

**Elevate scientific software development** through

- Availability of well-trained professionals
- Extensive training
- Education
- Community-building efforts
- Advocacy for both better software and for the people responsible for it

A modest **research component** on the "science of scientific software"

- In the spirit of SSSDU workshop report
- To help improve COLABS services and training and the broader community.

# The Research Software Engineer role

People with many job titles contribute to the development and maintenance of scientific software


Software professional


Computer Scientist


Domain scientist


Computational Scientist


and more…

Any of them can assume the role of research software engineer, full- or part-time

# The Research Software Engineer role

People with many job titles contribute to the development and maintenance of scientific software


Software professional


Computer Scientist


Domain scientist


Computational Scientist


and more…

Any of them can assume the role of research software engineer, full- or part-time

## COLABS use of the term “RSE”

- ..As part of a world-wide movement which has settled on the term "research software engineer" (RSE) for this role
  - "those who regularly use expertise in programming to advance research" -- US-RSE (professional society)
- Many research-oriented institutions have a hard time valuing contributions to software development and sustainment
- COLABS will advocate for better recognition and rewards for software-related contributions with member institutions and sponsors

# The Research Software Engineer role

People with many job titles contribute to the development and maintenance of scientific software


Software professional


Computer Scientist


Domain scientist


Computational Scientist


and more…

Any of them can assume the role of research software engineer, full- or part-time

## COLABS use of the term "RSE"

- ..As part of a world-wide movement which has settled on the term "research software engineer" (RSE) for this role
  - "those who regularly use expertise in programming to advance research" -- US-RSE (professional society)
- Many research-oriented institutions have a hard time valuing contributions to software development and sustainment
- COLABS will advocate for better recognition and rewards for software-related contributions with member institutions and sponsors

- Project staff can be paid by COLABS for a portion of their time (as agreed between individual, project, and COLABS)
- COLABS will also have staff recruited elsewhere
- Project staff in COLABS will have the option of joining the pool of service providers/subject matter experts

# Levels of Sustainability Services in COLABS

Sustainability focuses on **user-facing** aspects of the software (as distinct from R&D aspects)

### Assessment and Coaching

Help with assessment and improvement of software practices

- Leverage COLABS trainings to help clients in enhancing sustainability

### Essential Services

Baseline levels of effort

- Specific activities determined by project with advice from COLABS

### Central Services

Services that are more effectively centralized, e.g.,

- Setup and maintenance of CI/CD pipelines
- Building and maintaining containers
- etc.

### Advanced Services

Effort above baseline levels

- E.g., porting, major test development, refactoring, etc.
- Probably require a lightweight proposal process

### Subject Matter Experts

Pool of experts needed occasionally by projects, e.g.,

- Software architecture
- Test design
- UI/UX design
- Security protocols
- Porting to advanced hw
- etc.

# What Do We Mean By "Central Services"

- Some software engineering activities are hard for individual projects to implement effectively because...
  - The details are finicky and change frequently
  - Implementation in national lab/facility environments may impose additional requirements
  - Etc.
- CI/CD pipelines are an obvious example
  - Tools and best practices for implementation change frequently
  - Specialty environments (e.g., computing facilities, cloud services) and (more) secure environments (e.g., local hardware in national labs)
- Containerization is another example
  - Various tools, formats, limitations/restrictions for various flavors of containers
  - Different facilities, cloud services, etc. support different tools
  - Ensuring ability to effectively utilize specialize hardware (e.g., accelerators, interconnects)
- COLABS expects to have staff with high levels of knowledge and experience in such areas and can more efficiently support such needs for projects
  - Option to train project staff more deeply

# Additional COLABS Activities (1)

**Training**

- For client projects
  - Improve internal capabilities (“learning to fish”)
- For COLABS staff
  - Professional development – update and expand skills
- Also open to larger community

# Additional COLABS Activities (1)

## Training

- For client projects
  - Improve internal capabilities ("learning to fish")
- For COLABS staff
  - Professional development – update and expand skills
- Also open to larger community

## RSE Community of Practice

- Share experience and expertise
- Facilitate support of project needs
- Professional growth
- Engagement with larger community and advocacy for software professionals
  - E.g., US-RSE, international RSE organizations, etc.

# Additional COLABS Activities (2)

**Workforce development**

- We believe more people will be needed in RSE role – plan for significant workforce development activity to help address that need
  - Engage with teaching institutions to incorporate practical scientific software development and engineering into their curricula
  - Complement with internship and co-op programs to give students hands-on experience
- Emphasize inclusion of people from underrepresented and marginalized groups in education, internship, co-op efforts
  - Reach out to minority serving institutions and those that have no or limited connection to HPC
- COLABS staff will be encouraged to serve as instructors, mentors

# Additional COLABS Activities (2)

**Workforce development**

- We believe more people will be needed in RSE role – plan for significant workforce development activity to help address that need
  - Engage with teaching institutions to incorporate practical scientific software development and engineering into their curricula
  - Complement with internship and co-op programs to give students hands-on experience
- Emphasize inclusion of people from underrepresented and marginalized groups in education, internship, co-op efforts
  - Reach out to minority serving institutions and those that have no or limited connection to HPC
- COLABS staff will be encouraged to serve as instructors, mentors

**Research and development**

- A modest R&D activity with the goal of improving COLABS services and training
  - In the spirit of the findings of the Workshop on the Science of Scientific Software Development and Use (SSSDU) report (to be published soon)
- Involvement in software science R&D is another professional development opportunity for COLABS team members

Brochure for SSSDU available

# Working with COLABS

- **For projects with embedded "RSEs"**
  - COLABS can "adopt" them -- provide support, training, community, advocacy, and professional development opportunities
- **For projects needing sustainability services** (or perhaps unaware that they need them)
  - COLABS can provide needed services
- **Pool of subject matter experts available** to address occasional needs
  - Get the help without a dedicated hire

# Working with COLABS

- **For projects with embedded "RSEs"**
  - COLABS can "adopt" them -- provide support, training, community, advocacy, and professional development opportunities
- **For projects needing sustainability services** (or perhaps unaware that they need them)
  - COLABS can provide needed services
- **Pool of subject matter experts available** to address occasional needs
  - Get the help without a dedicated hire

- Goal 1: encourage **long-term relationships** between RSEs and projects
- Goal 2: **flexibility to projects** to tap others for some services
- Goal 3: **flexibility for RSEs** to engage with other projects

# Working with COLABS

- **For projects with embedded "RSEs"**
  - COLABS can "adopt" them -- provide support, training, community, advocacy, and professional development opportunities
- **For projects needing sustainability services** (or perhaps unaware that they need them)
  - COLABS can provide needed services
- **Pool of subject matter experts available** to address occasional needs
  - Get the help without a dedicated hire

- Goal 1: encourage **long-term relationships** between RSEs and projects
- Goal 2: **flexibility to projects** to tap others for some services
- Goal 3: **flexibility for RSEs** to engage with other projects

**Benefits to COLABS Members**

- Professional development and growth opportunities
- Training to stay current with the state-of-the-art
- Access to a community of like-minded people
- Appropriate recognition of contributions to COLABS clients
- Advocacy for recognition and career path
- Greater (funding) stability for staff in RSE roles

# Working with COLABS

- **For projects with embedded "RSEs"**
  - COLABS can "adopt" them -- provide support, training, community, advocacy, and professional development opportunities
- **For projects needing sustainability services** (or perhaps unaware that they need them)
  - COLABS can provide needed services
- **Pool of subject matter experts available** to address occasional needs
  - Get the help without a dedicated hire

- Goal 1: encourage **long-term relationships** between RSEs and projects
- Goal 2: **flexibility to projects** to tap others for some services
- Goal 3: **flexibility for RSEs** to engage with other projects

**Benefits to COLABS Members**

- Professional development and growth opportunities
- Training to stay current with the state-of-the-art
- Access to a community of like-minded people
- Appropriate recognition of contributions to COLABS clients
- Advocacy for recognition and career path
- Greater (funding) stability for staff in RSE roles

**Benefits to Project (PI)**

- Greater stability for services critical for project
  - Less need to cut corners on quality and user-engagement
  - Greater continuity in face of erratic funding cycles
- Support for capabilities that can be hard to sustain (especially for smaller teams)
- More diverse expertise available when needed

# Working with COLABS -- Special Cases

## Incubation

- Help with initial setup and software process development
- Training project personnel
- Help bring the project to a self-sufficient operational level

## Adoption

- A project from outside the normal constituency
  - But important to a client project, sponsor, or the larger community
  - E.g., a third-party dependency in need of additional support for sustainability or tiding-over until a replacement can be found
  - Probably triggered by a COLABS client or sponsor request
- Would require agreement of sponsor, and adopted project (of course)
- Note: in most cases need participation of knowledgeable people to do much beyond "life support"
  - Contributors to adopted project, maybe advanced users

# Benefits of COLABS for Sponsors

- Increased attention to, and action on software sustainability
- Economies of scale in sustainability services
  - The design of COLABS can support the needs of multiple sponsors
  - ASCR is the initial sponsor
  - Can imagine other DOE offices funding COLABS to support application projects, SciDAC projects, experimental and observational user facilities, etc.
- Better match of tasks and expertise for more efficiency
- Highly trained workforce
- More professional development, career support → greater retention
  - COLABS also plans a significant workforce development component to build a stronger, deeper pipeline for growing needs for people in the RSE role
- Possibility of bridging funding gaps for useful software

# Still Under Discussion

## Allocation of COLABS Resources

- Training available to all, to promote better software practices
- Sponsors engagement in defining a "constituency" of projects that we're "allowed" to support at all levels
- Coaching available to all projects in the constituency, to help implement better software practices
- Essential and central services widely available within the constituency
- Allocation for higher service levels is under discussion
- Periodic review of portfolio with Advisory Board to ensure value for stakeholders (projects, users, sponsors)
- Projects can buy additional services directly from COLABS, if desired

# Still Under Discussion

## Allocation of COLABS Resources

- Training available to all, to promote better software practices
- Sponsors engagement in defining a "constituency" of projects that we're "allowed" to support at all levels
- Coaching available to all projects in the constituency, to help implement better software practices
- Essential and central services widely available within the constituency
- Allocation for higher service levels is under discussion
- Periodic review of portfolio with Advisory Board to ensure value for stakeholders (projects, users, sponsors)
- Projects can buy additional services directly from COLABS, if desired

## Engagement with Industry, Other Sponsors and Projects

- Companies and other agencies could buy into COLABS as sponsors
- COLABS can support industry/other agency projects with sponsor concurrence
  - e.g., DOE currently engages with other agencies as a leading federal agency in HPC/scientific computing; could do the same for scientific software stewardship
- COLABS can collaborate with other software sustainability organizations
  - Potentially, federation of SSOs

# COLABS Design is Scalable in Multiple Dimensions

**Organizational scalability**

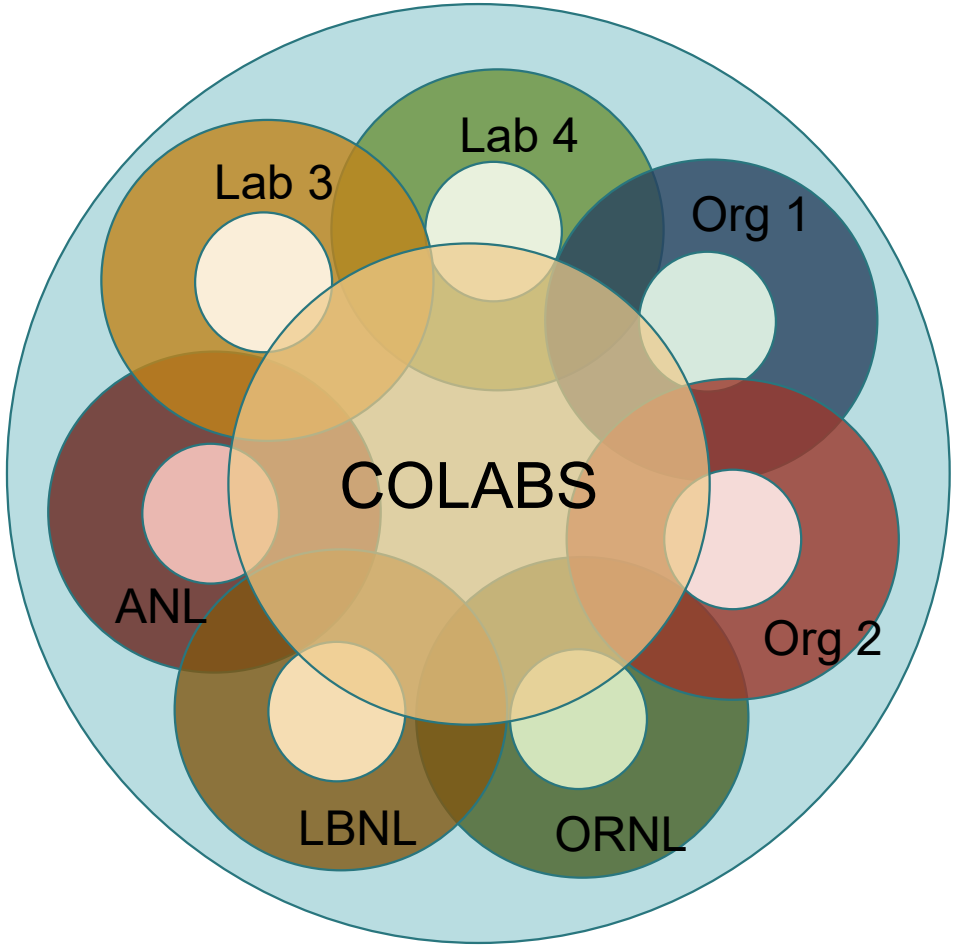
- We envision an institutional footprint for COLABS which is similar to that of the constituency it is meant to serve
- Tempered by the desire to maintain critical mass for communities of practice to be meaningful at the institutional level

**Budgetary scalability**

- The services COLABS provides can be scaled to the budget available (within limits, of course) -- more funding → more services

**Scalability to additional sponsors**

- The same structure and governance model, and to a large extent the same (types of) people can support multiple sponsors
- ASCR is first sponsor. Others could be Office of Science offices, user facilities, DOE applied offices, etc.



# Summary

- **COLABS is a software sustainability organization designed to help elevate, sustain, and steward scientific software**
  - Better software quality, reliability, sustainability
  - Better knowledge and skills for software development and maintenance
  - Better recognition and support for the people who do the work
  - Pipelines to develop the growing scientific software workforce DOE will need
- **Services to be available as broadly as possible**
- **Growing the RSE community of practice**
  - Flexibility for COLABS staff to work with different projects, work as SMEs
  - Sharing knowledge and skills with others, learning from them
  - Opportunities for training, education, mentoring on RSE skills & experience

**Interested? Talk to Us!**

- **Project leaders**
  - What services would you like to see COLABS provide?
  - What levels of effort do you need to sustain your software?
- **Prospective COLABS members**
  - What would you find attractive about being a member of COLABS?
  - What skills do you have, and what would you like to learn?

***Science through computing is, at best, as good as the software that produces it***