How Open Source Supports the Largest Computers on the Planet

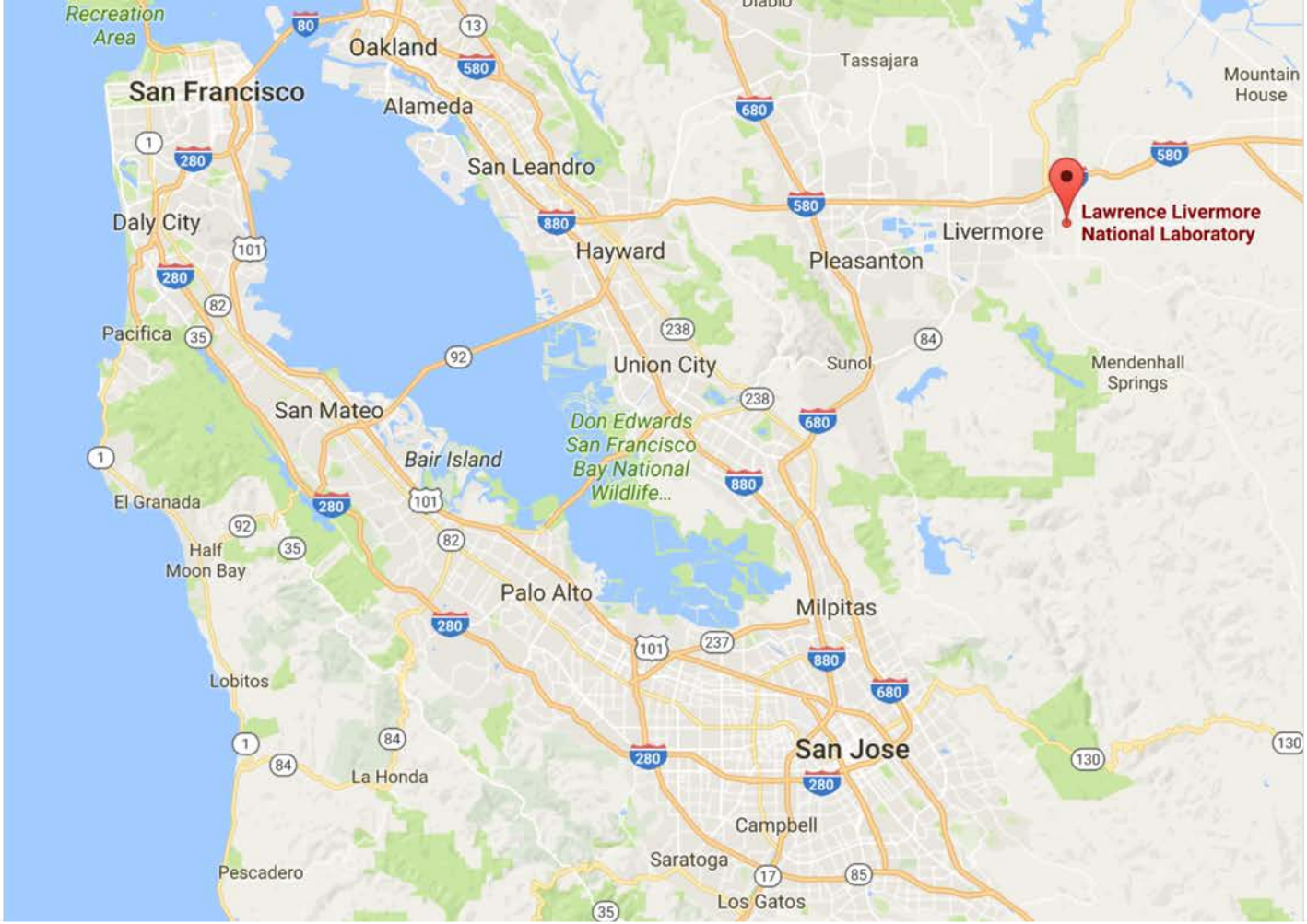
Best Practices for HPC Software Developers

Ian Lee
Lawrence Livermore National Laboratory

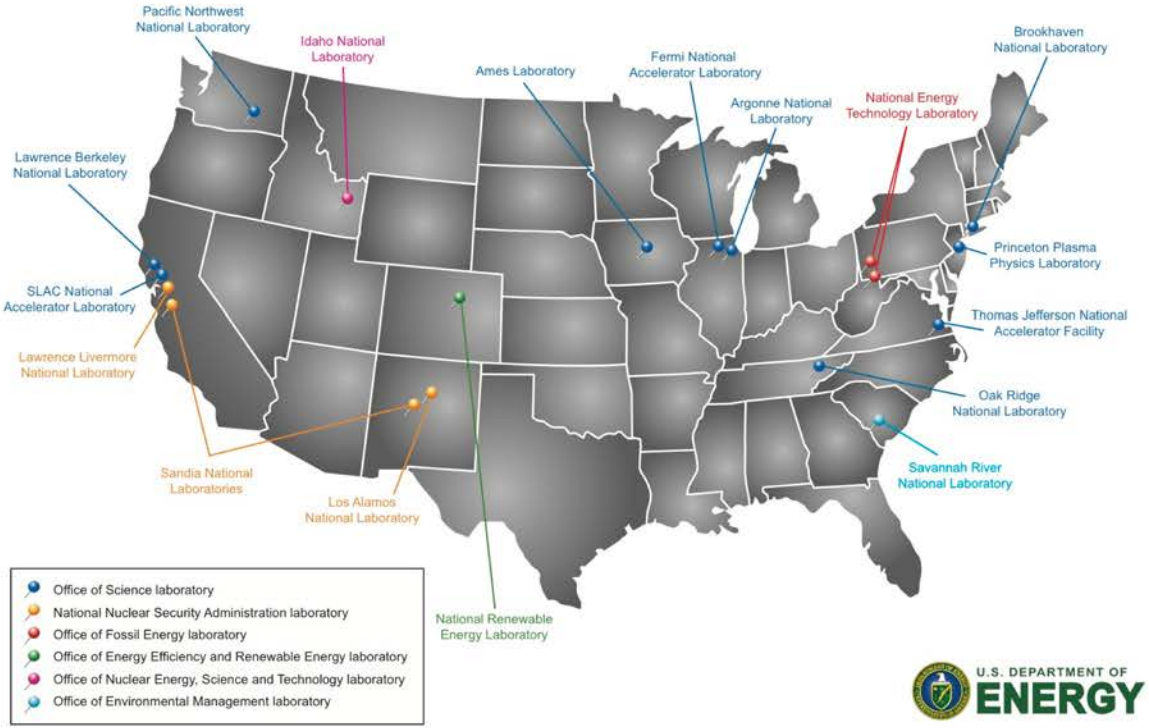
July 18, 2018

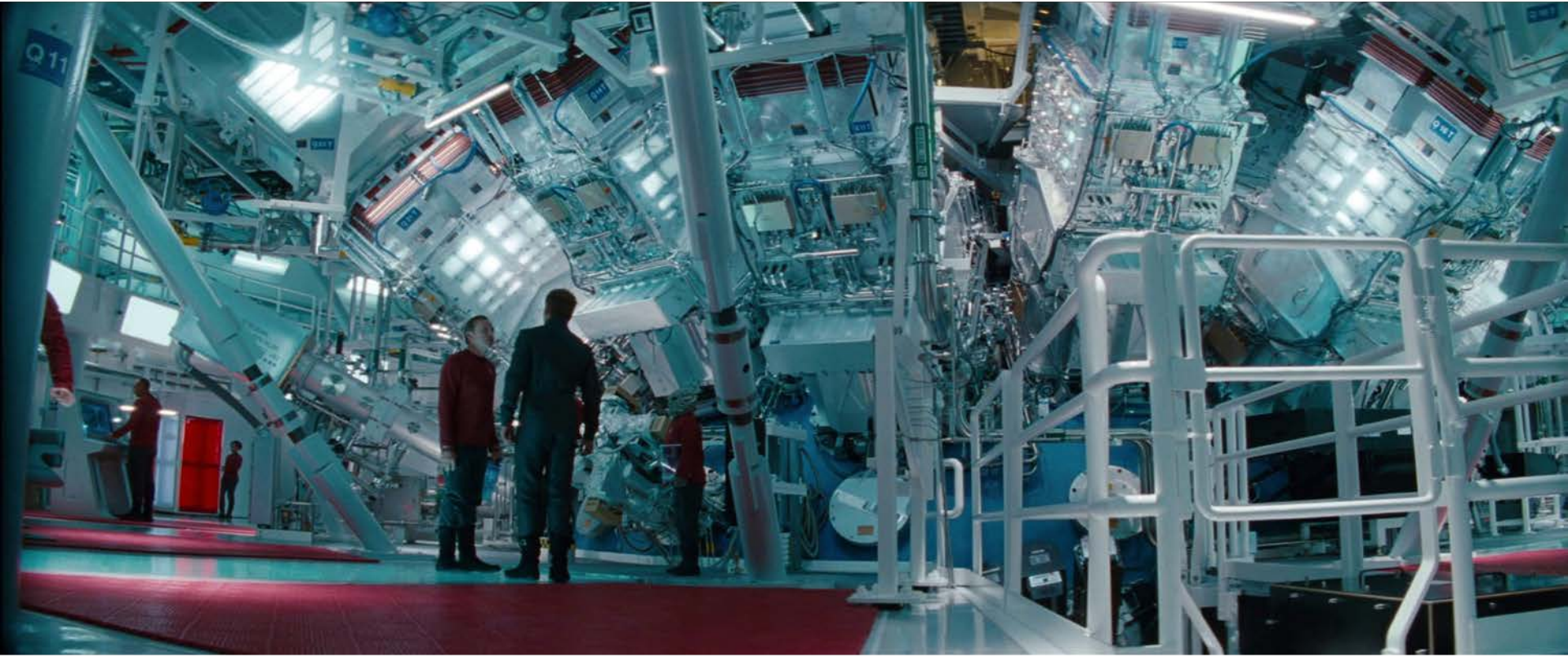
LLNL-PRES-754800
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

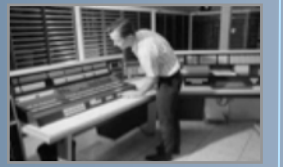

Recreation Area
Oakland
San Francisco
Alameda
San Leandro
Tassajara
Mountain House
Daly City
Hayward
Livermore
Lawrence Livermore National Laboratory
Pleasanton
Pacifica
Union City
Sunol
Mendenhall Springs
San Mateo
Don Edwards San Francisco Bay National Wildlife...
Bair Island
El Granada
Half Moon Bay
Palo Alto
Milpitas
Lobitos
San Jose
La Honda
Campbell
Saratoga
Pescadero
Los Gatos

Department of Energy National Laboratories

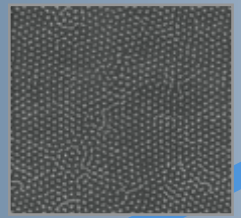



Q11

https://pixabay.com/get/e833b10d2af4083ed1534705fb0938c9bd22ffd41db612439df7c17ba0/silos-1602209_1920.jpg

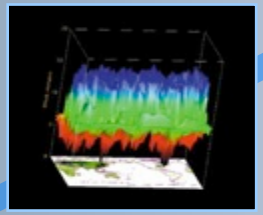


CDC 3600

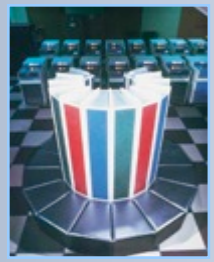


Pioneering simulations of particle tracking

CDC 7600



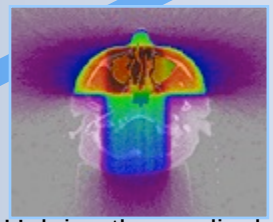
Ozone mixing models



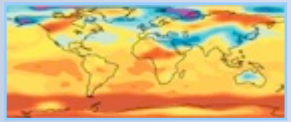
CRAY 1

Dynamics in three dimensions

ASCI Blue-Pacific

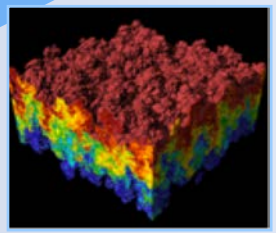


Helping the medical community plan radiation treatment



Global climate modeling

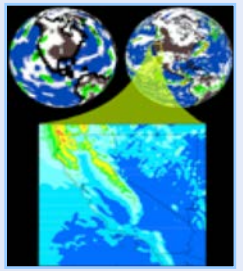
BlueGene



Breakthrough visualizations of mixing fluids

Unprecedented dislocation dynamics simulations

Petascale and exascale computing



Detailed predictions of ecosystems

Top500.org

- 3 out of 16 #1 systems over last 20 years

ASCI White
Nov 2000 – Nov 2001

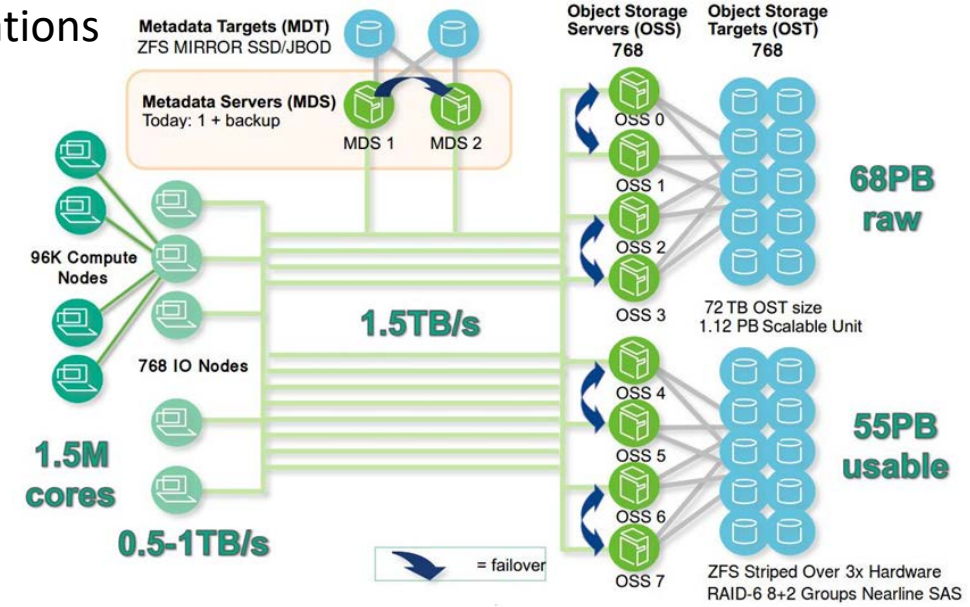
BlueGene/L
Nov 2004 – Nov 2007

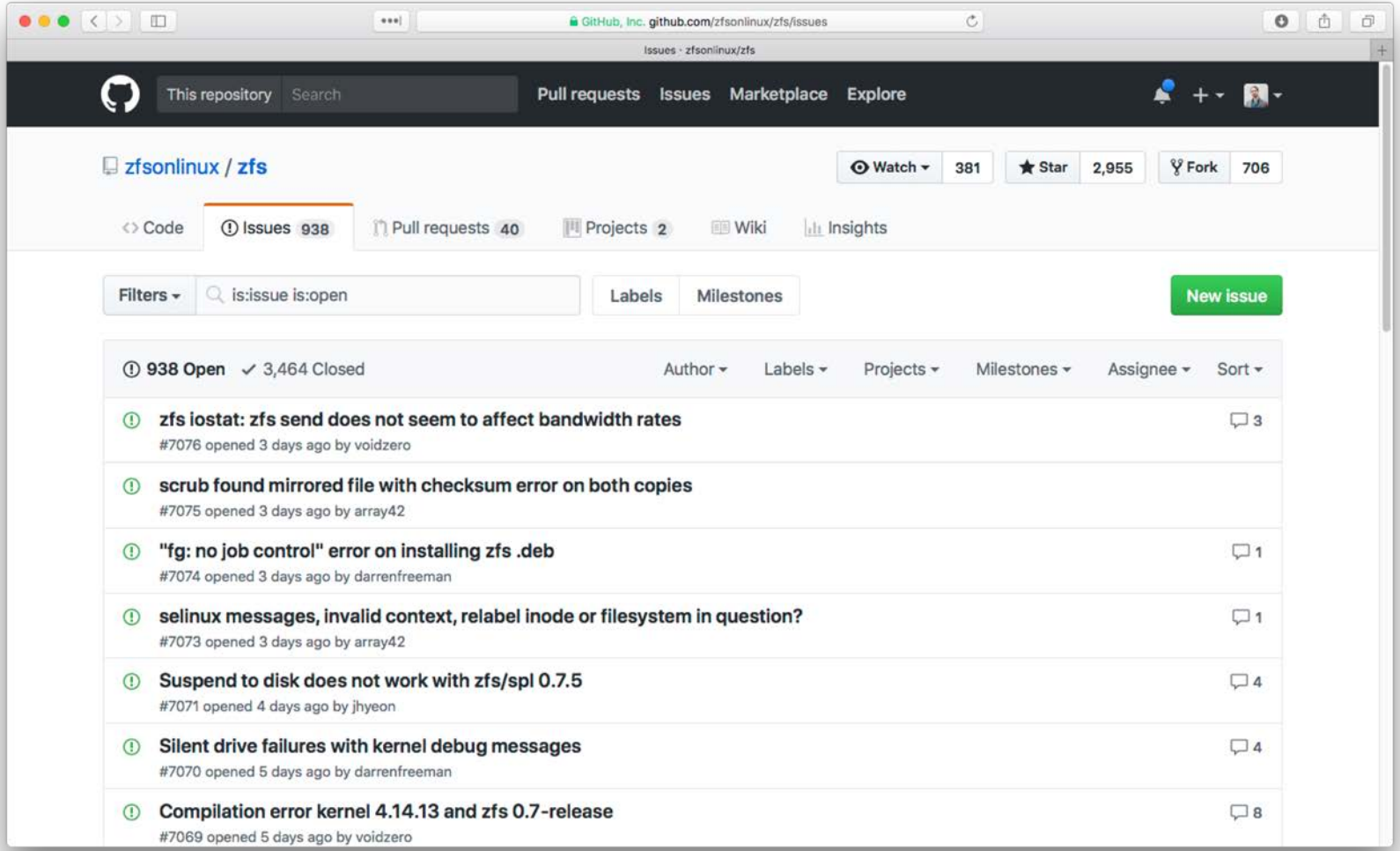
Sequoia
June 2012

Sierra

ZFS on Linux

- ZFS is an open source filesystem and volume manager designed to address the limitations of existing storage solutions
- 2011: Available for Linux
- Ten LLNL filesystems, totaling ~ 100PB
- Ships in Ubuntu 16.04




GitHub, Inc. github.com/zfsonlinux/zfs/issues
Issues · zfsonlinux/zfs
This repository
Search
Pull requests
Issues
Marketplace
Explore
zfsonlinux / zfs
Watch 381
Star 2,955
Fork 706
Code
Issues 938
Pull requests 40
Projects 2
Wiki
Insights
Filters
is:issue is:open
Labels
Milestones
New issue
938 Open
3,464 Closed
Author
Labels
Projects
Milestones
Assignee
Sort
zfs iostat: zfs send does not seem to affect bandwidth rates
#7076 opened 3 days ago by voidzero
3
scrub found mirrored file with checksum error on both copies
#7075 opened 3 days ago by array42
"fg: no job control" error on installing zfs .deb
#7074 opened 3 days ago by darrenfreeman
1
selinux messages, invalid context, relabel inode or filesystem in question?
#7073 opened 3 days ago by array42
1
Suspend to disk does not work with zfs/spl 0.7.5
#7071 opened 4 days ago by jhyeon
4
Silent drive failures with kernel debug messages
#7070 opened 5 days ago by darrenfreeman
4
Compilation error kernel 4.14.13 and zfs 0.7-release
#7069 opened 5 days ago by voidzero
8

GitHub, Inc. github.com/zfsonlinux/zfs/blob/master/.github/CONTRIBUTING.md

zfs/CONTRIBUTING.md at master · zfsonlinux/zfs

This repository Search Pull requests Issues Marketplace Explore

zfsonlinux / zfs

Watch 381 Star 2,956 Fork 706

Code Issues 938 Pull requests 40 Projects 2 Wiki Insights

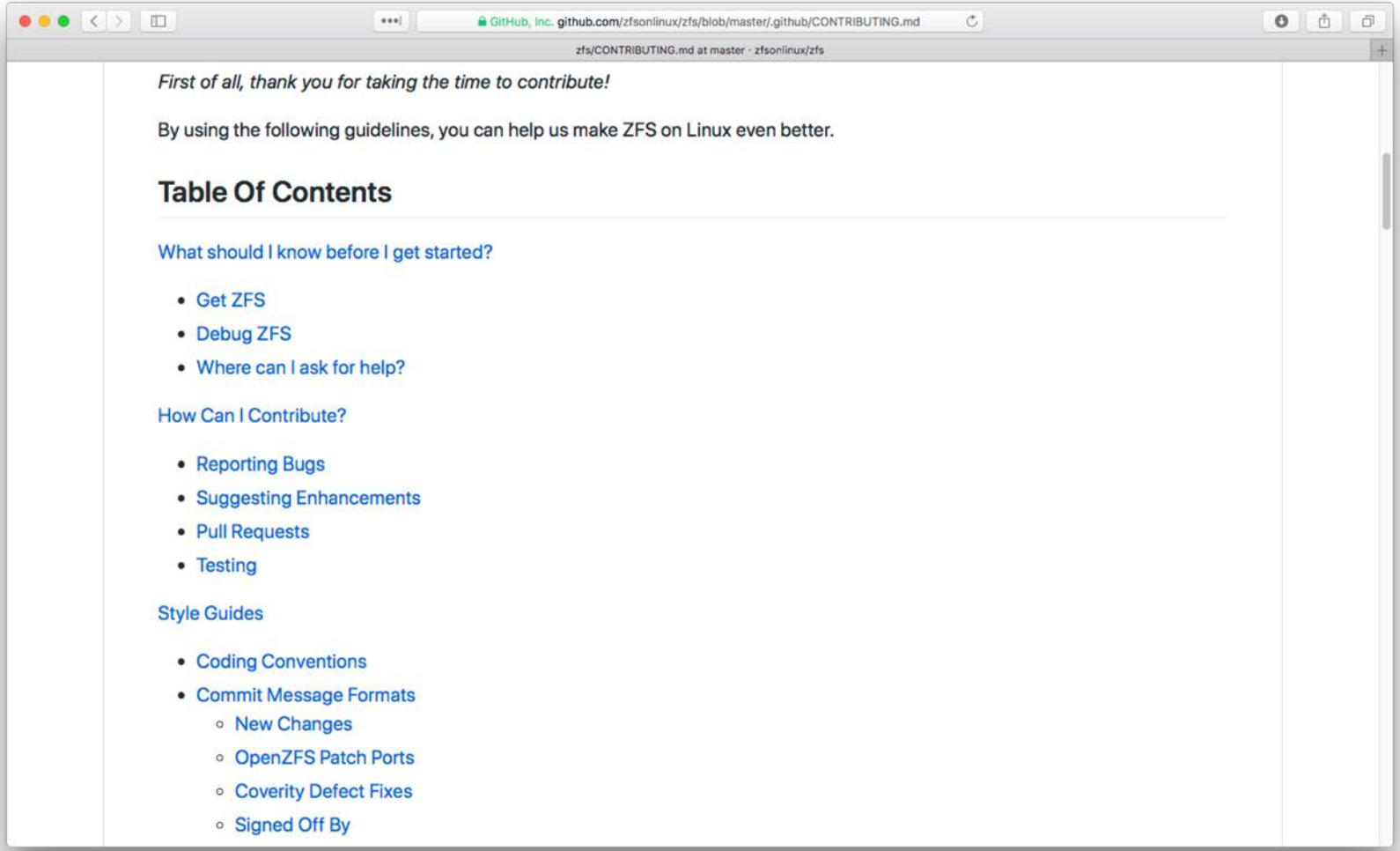
Branch: master zfs / .github / CONTRIBUTING.md Find file Copy path

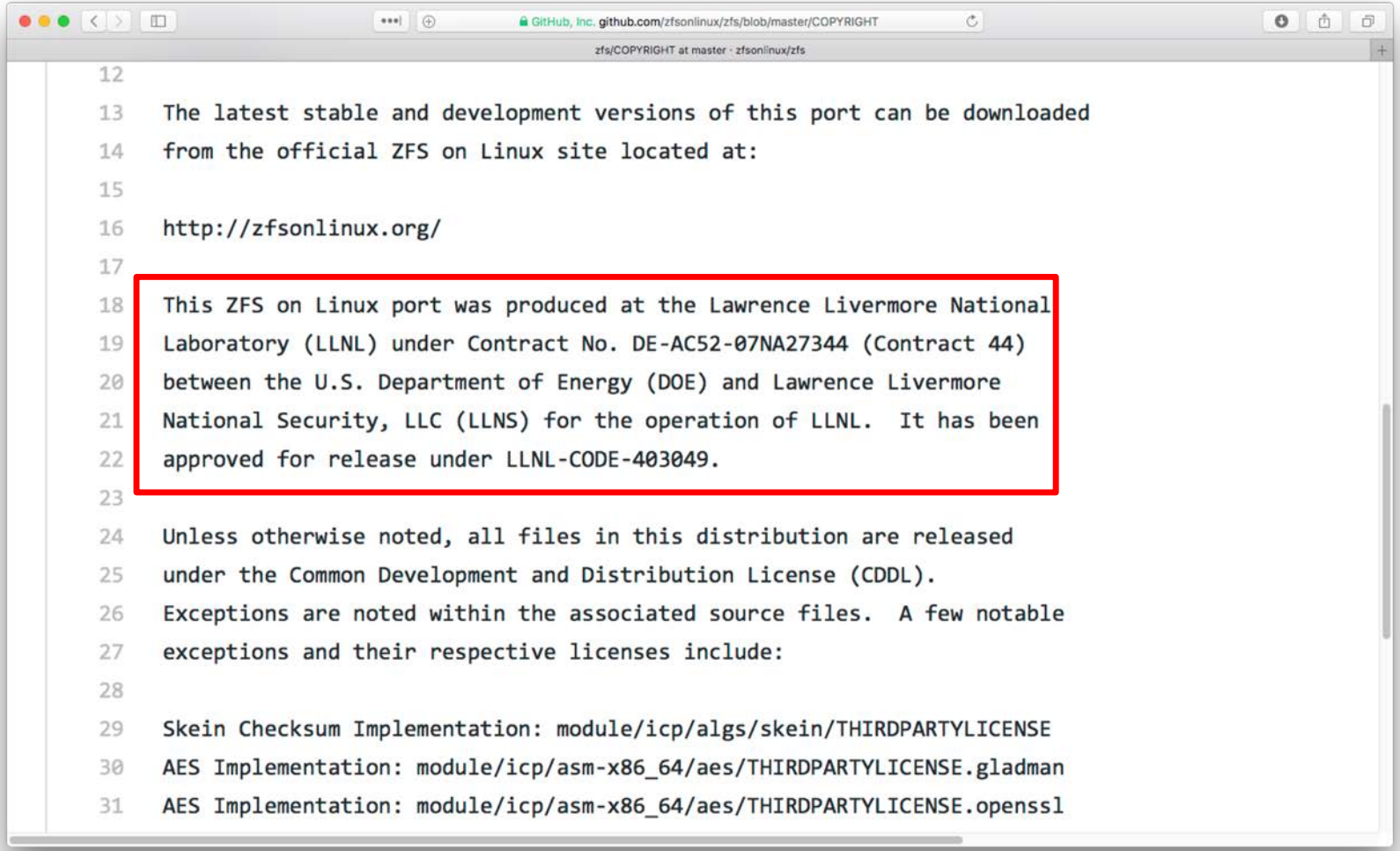
dinatale2 commitcheck: Multiple OpenZFS ports in commit 69b229b on Oct 26, 2017

3 contributors

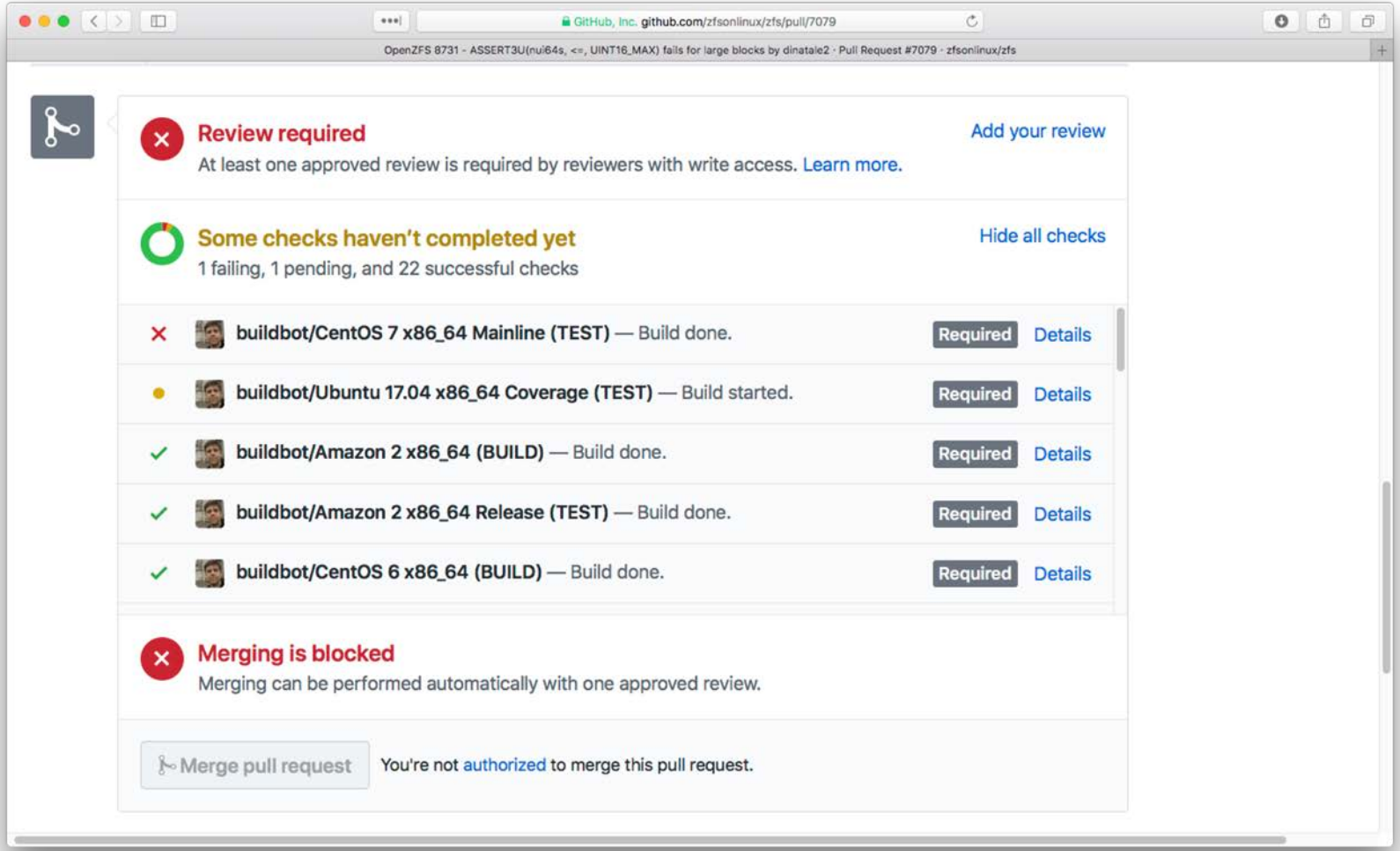
291 lines (230 sloc) 11.6 KB Raw Blame History

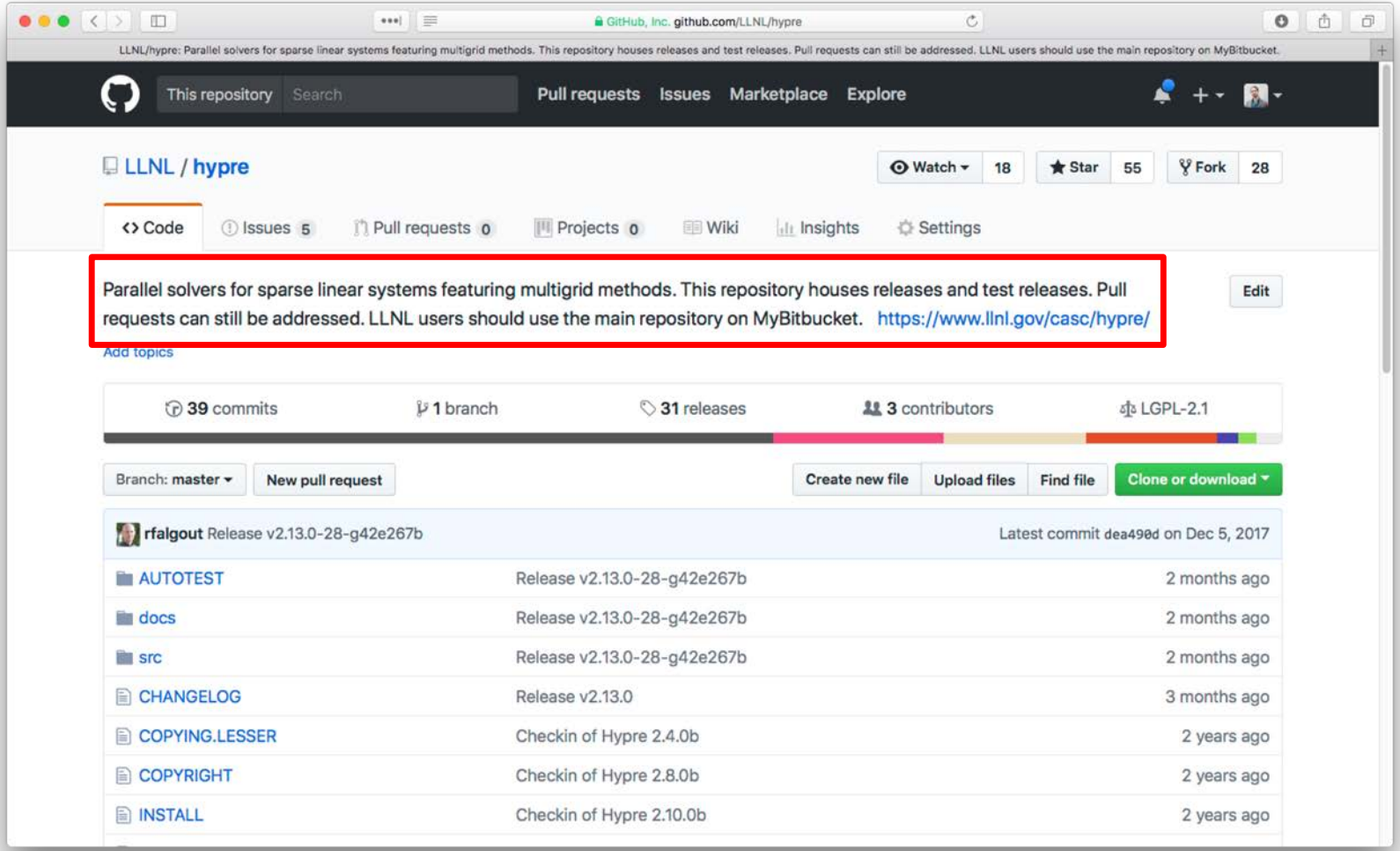
Contributing to ZFS on Linux

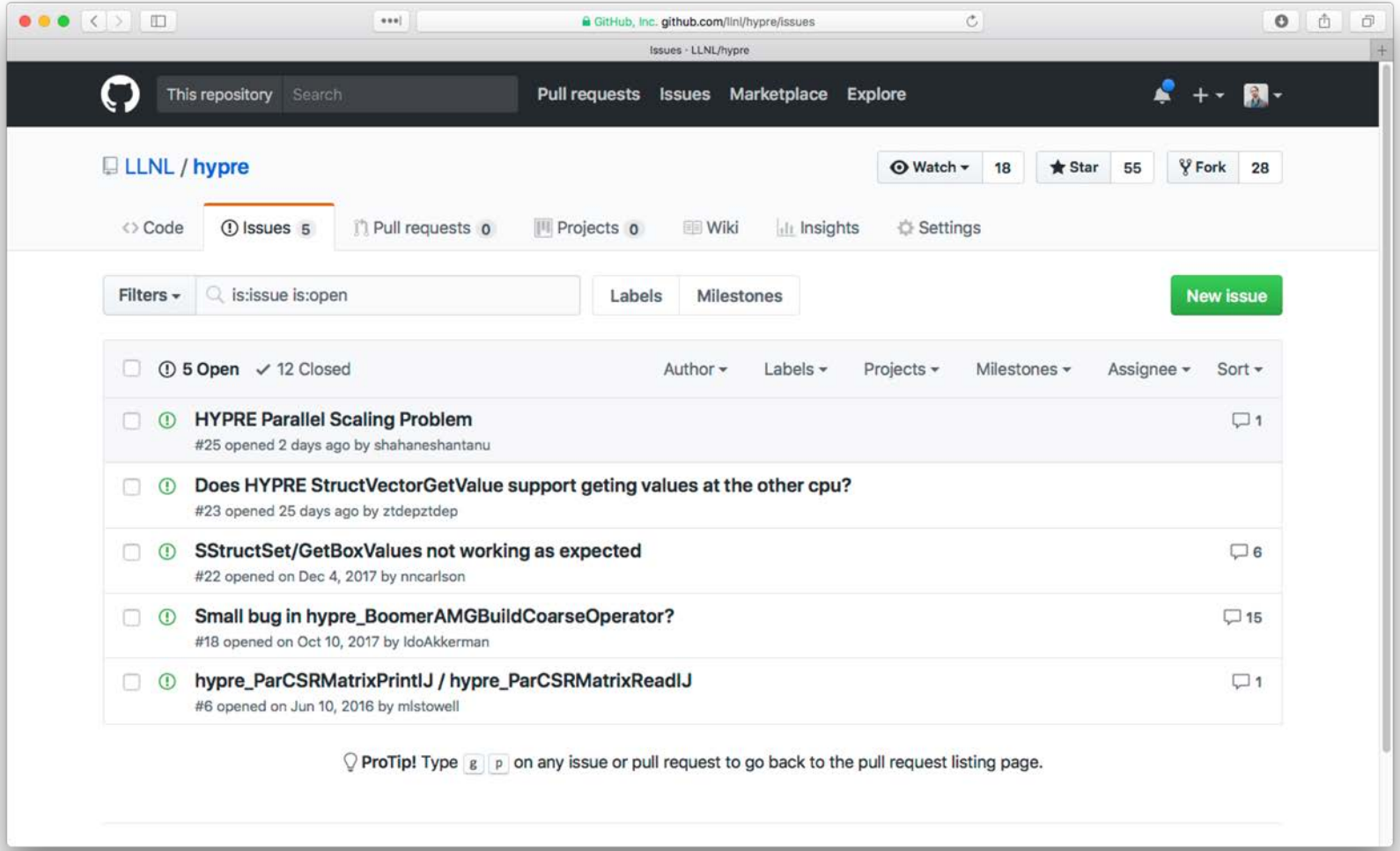

GitHub, Inc. github.com/zfsonlinux/zfs/blob/master/.github/CONTRIBUTING.md
zfs/CONTRIBUTING.md at master · zfsonlinux/zfs
First of all, thank you for taking the time to contribute!
By using the following guidelines, you can help us make ZFS on Linux even better.
Table Of Contents
What should I know before I get started?
Get ZFS
Debug ZFS
Where can I ask for help?
How Can I Contribute?
Reporting Bugs
Suggesting Enhancements
Pull Requests
Testing
Style Guides
Coding Conventions
Commit Message Formats
New Changes
OpenZFS Patch Ports
Coverity Defect Fixes
Signed Off By

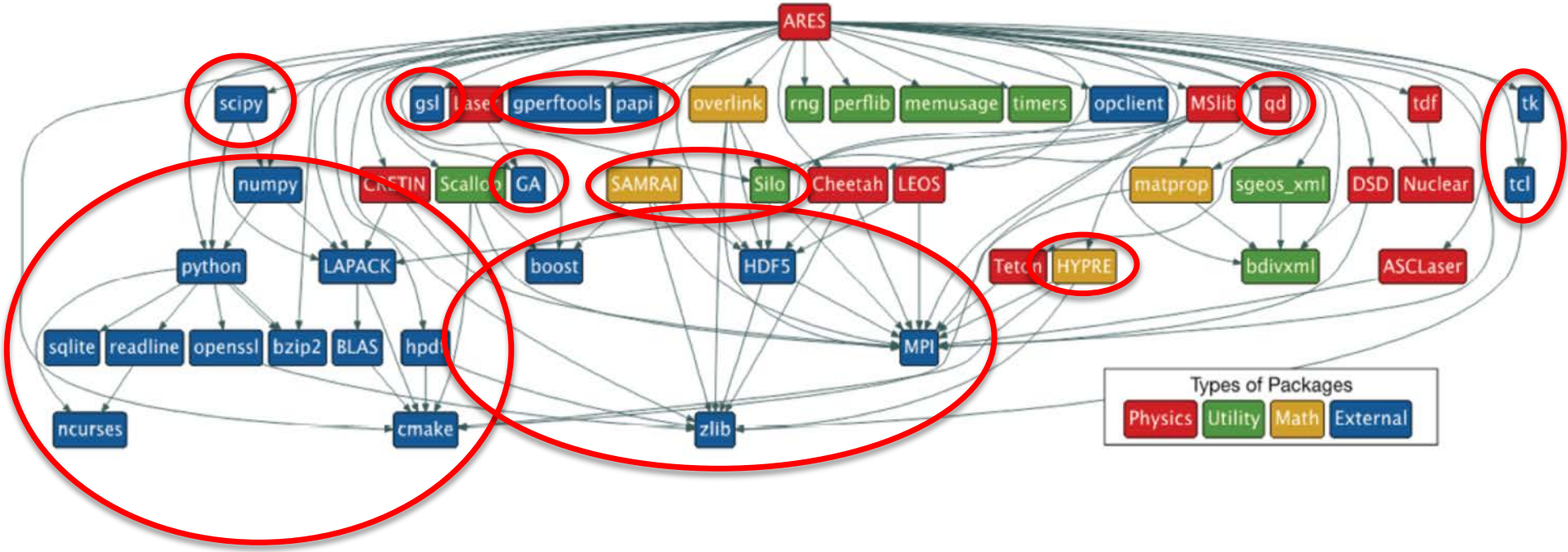


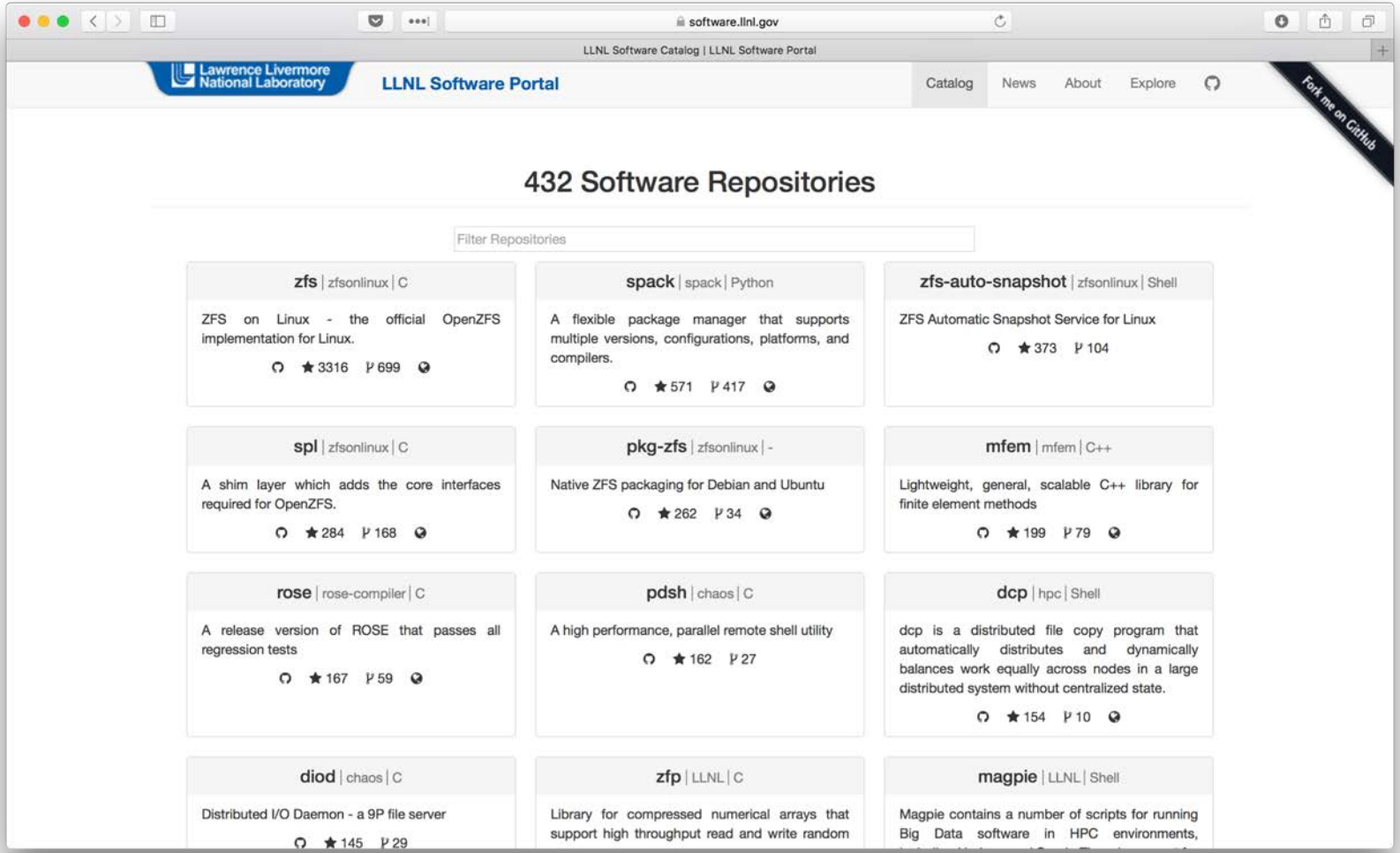

```
12
13  The latest stable and development versions of this port can be downloaded
14  from the official ZFS on Linux site located at:
15
16  http://zfsonlinux.org/
17
18  This ZFS on Linux port was produced at the Lawrence Livermore National
19  Laboratory (LLNL) under Contract No. DE-AC52-07NA27344 (Contract 44)
20  between the U.S. Department of Energy (DOE) and Lawrence Livermore
21  National Security, LLC (LLNS) for the operation of LLNL.  It has been
22  approved for release under LLNL-CODE-403049.
23
24  Unless otherwise noted, all files in this distribution are released
25  under the Common Development and Distribution License (CDDL).
26  Exceptions are noted within the associated source files.  A few notable
27  exceptions and their respective licenses include:
28
29  Skein Checksum Implementation: module/icp/algs/skein/THIRDPARTYLICENSE
30  AES Implementation: module/icp/asm-x86_64/aes/THIRDPARTYLICENSE.gladman
31  AES Implementation: module/icp/asm-x86_64/aes/THIRDPARTYLICENSE.openssl
```


GitHub, Inc. github.com/zfsonlinux/zfs/pull/7079
OpenZFS 8731 - ASSERT3U(nui64s, <=, UINT16_MAX) fails for large blocks by dinatale2 · Pull Request #7079 · zfsonlinux/zfs
Review required
At least one approved review is required by reviewers with write access. Learn more.
Add your review
Some checks haven't completed yet
1 failing, 1 pending, and 22 successful checks
Hide all checks
buildbot/CentOS 7 x86_64 Mainline (TEST) — Build done.
Required
Details
buildbot/Ubuntu 17.04 x86_64 Coverage (TEST) — Build started.
Required
Details
buildbot/Amazon 2 x86_64 (BUILD) — Build done.
Required
Details
buildbot/Amazon 2 x86_64 Release (TEST) — Build done.
Required
Details
buildbot/CentOS 6 x86_64 (BUILD) — Build done.
Required
Details
Merging is blocked
Merging can be performed automatically with one approved review.
Merge pull request
You're not authorized to merge this pull request.

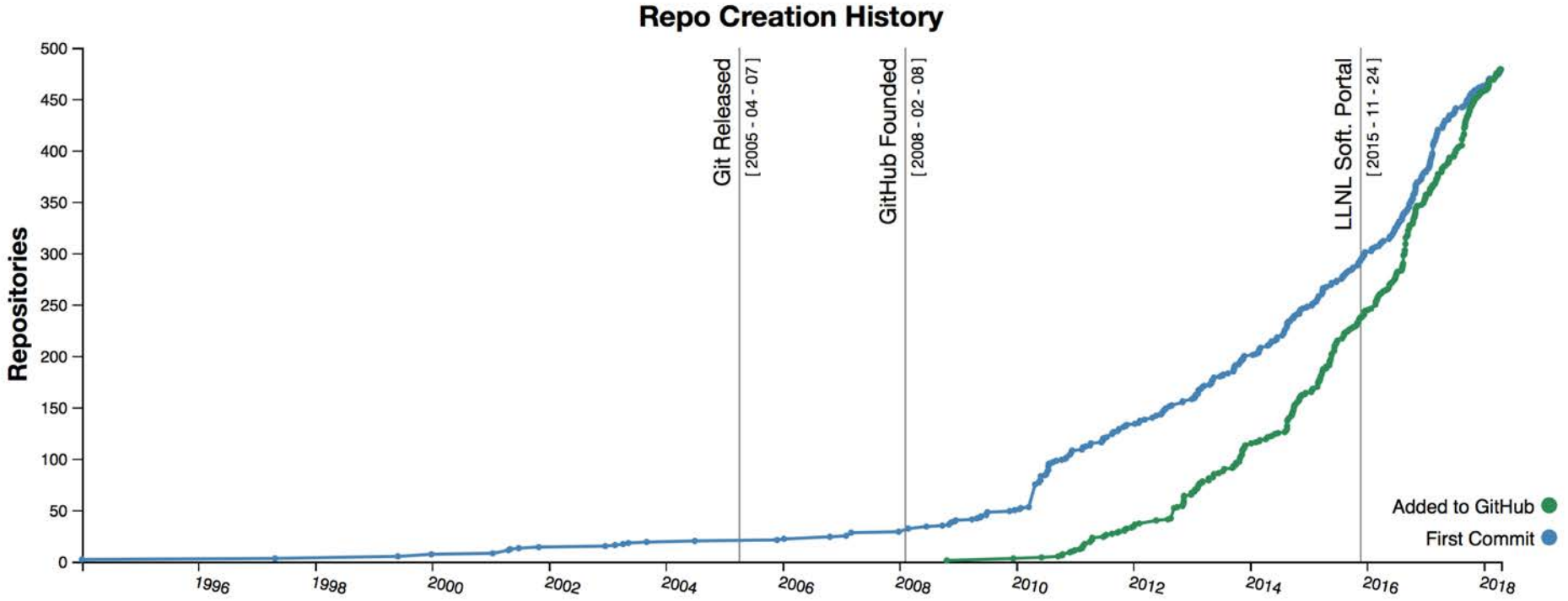

GitHub, Inc. github.com/LLNL/hypre
LLNL/hypre: Parallel solvers for sparse linear systems featuring multigrid methods. This repository houses releases and test releases. Pull requests can still be addressed. LLNL users should use the main repository on MyBitbucket.
This repository
Search
Pull requests
Issues
Marketplace
Explore
LLNL / hypre
Watch 18
Star 55
Fork 28
Code
Issues 5
Pull requests 0
Projects 0
Wiki
Insights
Settings
Parallel solvers for sparse linear systems featuring multigrid methods. This repository houses releases and test releases. Pull requests can still be addressed. LLNL users should use the main repository on MyBitbucket. https://www.llnl.gov/casc/hypre/
Edit
Add topics
39 commits
1 branch
31 releases
3 contributors
LGPL-2.1
Branch: master
New pull request
Create new file
Upload files
Find file
Clone or download
rfalgout Release v2.13.0-28-g42e267b
Latest commit dea490d on Dec 5, 2017
AUTOTEST Release v2.13.0-28-g42e267b 2 months ago
docs Release v2.13.0-28-g42e267b 2 months ago
src Release v2.13.0-28-g42e267b 2 months ago
CHANGELOG Release v2.13.0 3 months ago
COPYING.LESSER Checkin of Hypre 2.4.0b 2 years ago
COPYRIGHT Checkin of Hypre 2.8.0b 2 years ago
INSTALL Checkin of Hypre 2.10.0b 2 years ago


GitHub, Inc. github.com/llnl/hypre/issues
Issues · LLNL/hypre
This repository
Search
Pull requests
Issues
Marketplace
Explore
LLNL / hypre
Watch 18
Star 55
Fork 28
Code
Issues 5
Pull requests 0
Projects 0
Wiki
Insights
Settings
Filters
is:issue is:open
Labels
Milestones
New issue
5 Open
12 Closed
Author
Labels
Projects
Milestones
Assignee
Sort
HYPRE Parallel Scaling Problem
#25 opened 2 days ago by shahaneshantanu
1
Does HYPRE StructVectorGetValue support geting values at the other cpu?
#23 opened 25 days ago by ztdepztdep
SStructSet/GetBoxValues not working as expected
#22 opened on Dec 4, 2017 by nncarlson
6
Small bug in hypre_BoomerAMGBuildCoarseOperator?
#18 opened on Oct 10, 2017 by IdoAkkerman
15
hypre_ParCSRMatrixPrintIJ / hypre_ParCSRMatrixReadIJ
#6 opened on Jun 10, 2016 by mlstowell
1
ProTip! Type g p on any issue or pull request to go back to the pull request listing page.

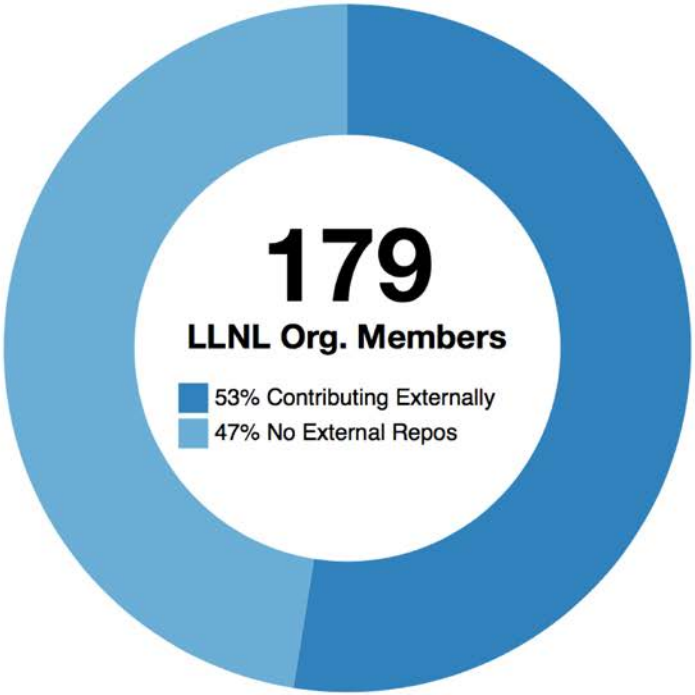

ARES
scipy
gsl
gperftools
papi
overlink
rng
perflib
memusage
timers
opclient
MSlib
qd
tdf
tk
numpy
SAMRAI
Silo
Cheetah
LEOS
matprop
sgeos_xml
DSD
Nuclear
tcl
python
LAPACK
boost
HDF5
HYPRE
bdivxml
ASCLaser
sqlite
readline
openssl
bzip2
BLAS
MPI
ncurses
cmake
zlib
Types of Packages
Physics
Utility
Math
External

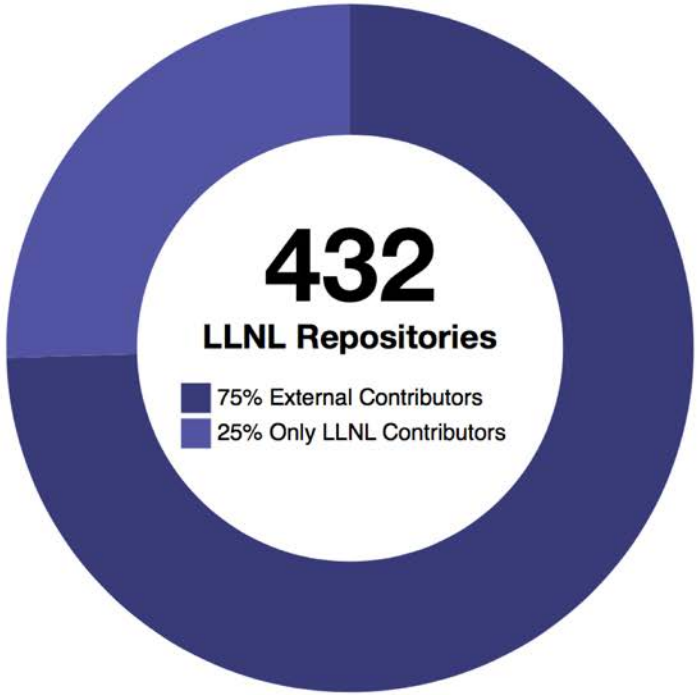

software.llnl.gov
LLNL Software Catalog | LLNL Software Portal
Lawrence Livermore National Laboratory
LLNL Software Portal
Catalog
News
About
Explore
Fork me on GitHub
432 Software Repositories
Filter Repositories
zfs | zfsonlinux | C
ZFS on Linux - the official OpenZFS implementation for Linux.
★ 3316 ⑂ 699
spack | spack | Python
A flexible package manager that supports multiple versions, configurations, platforms, and compilers.
★ 571 ⑂ 417
zfs-auto-snapshot | zfsonlinux | Shell
ZFS Automatic Snapshot Service for Linux
★ 373 ⑂ 104
spl | zfsonlinux | C
A shim layer which adds the core interfaces required for OpenZFS.
★ 284 ⑂ 168
pkg-zfs | zfsonlinux | -
Native ZFS packaging for Debian and Ubuntu
★ 262 ⑂ 34
mfem | mfem | C++
Lightweight, general, scalable C++ library for finite element methods
★ 199 ⑂ 79
rose | rose-compiler | C
A release version of ROSE that passes all regression tests
★ 167 ⑂ 59
pdsh | chaos | C
A high performance, parallel remote shell utility
★ 162 ⑂ 27
dcp | hpc | Shell
dcp is a distributed file copy program that automatically distributes and dynamically balances work equally across nodes in a large distributed system without centralized state.
★ 154 ⑂ 10
diod | chaos | C
Distributed I/O Daemon - a 9P file server
★ 145 ⑂ 29
zfp | LLNL | C
Library for compressed numerical arrays that support high throughput read and write random
magpie | LLNL | Shell
Magpie contains a number of scripts for running Big Data software in HPC environments,

LLNL Open Source Presence

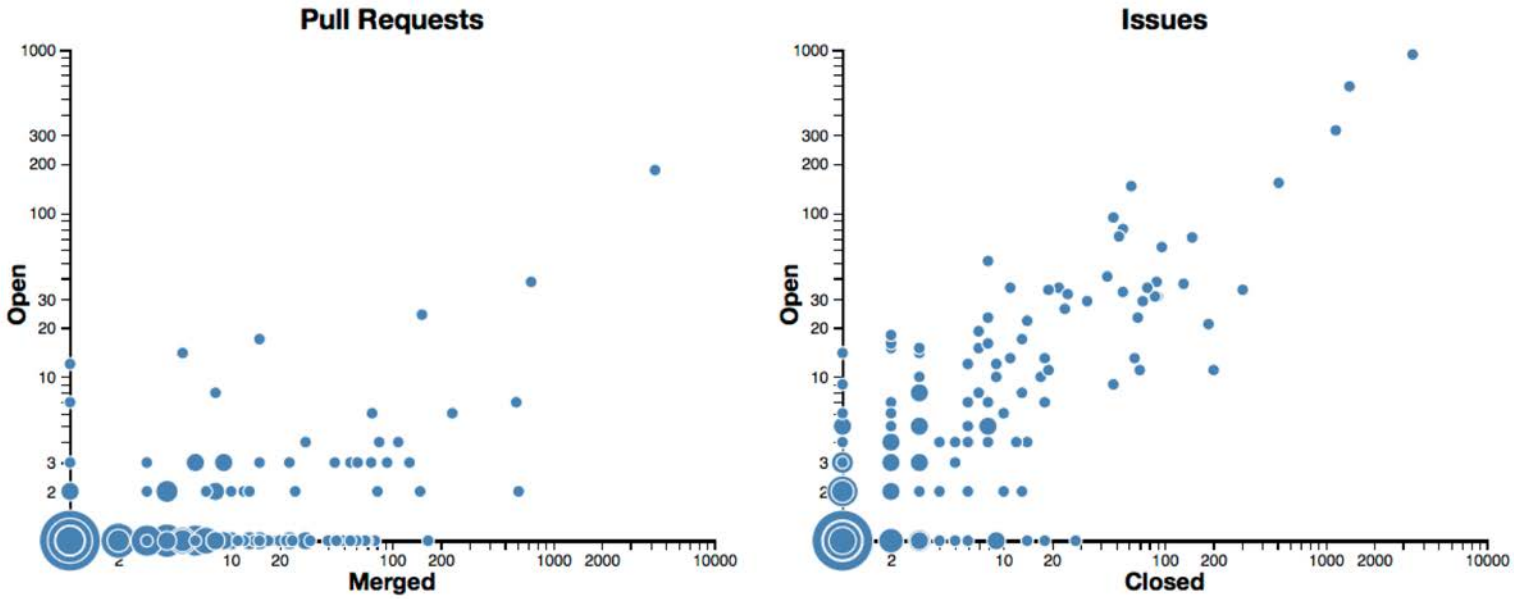


LLNL Open Source Engagement





LLNL Open Source Activities

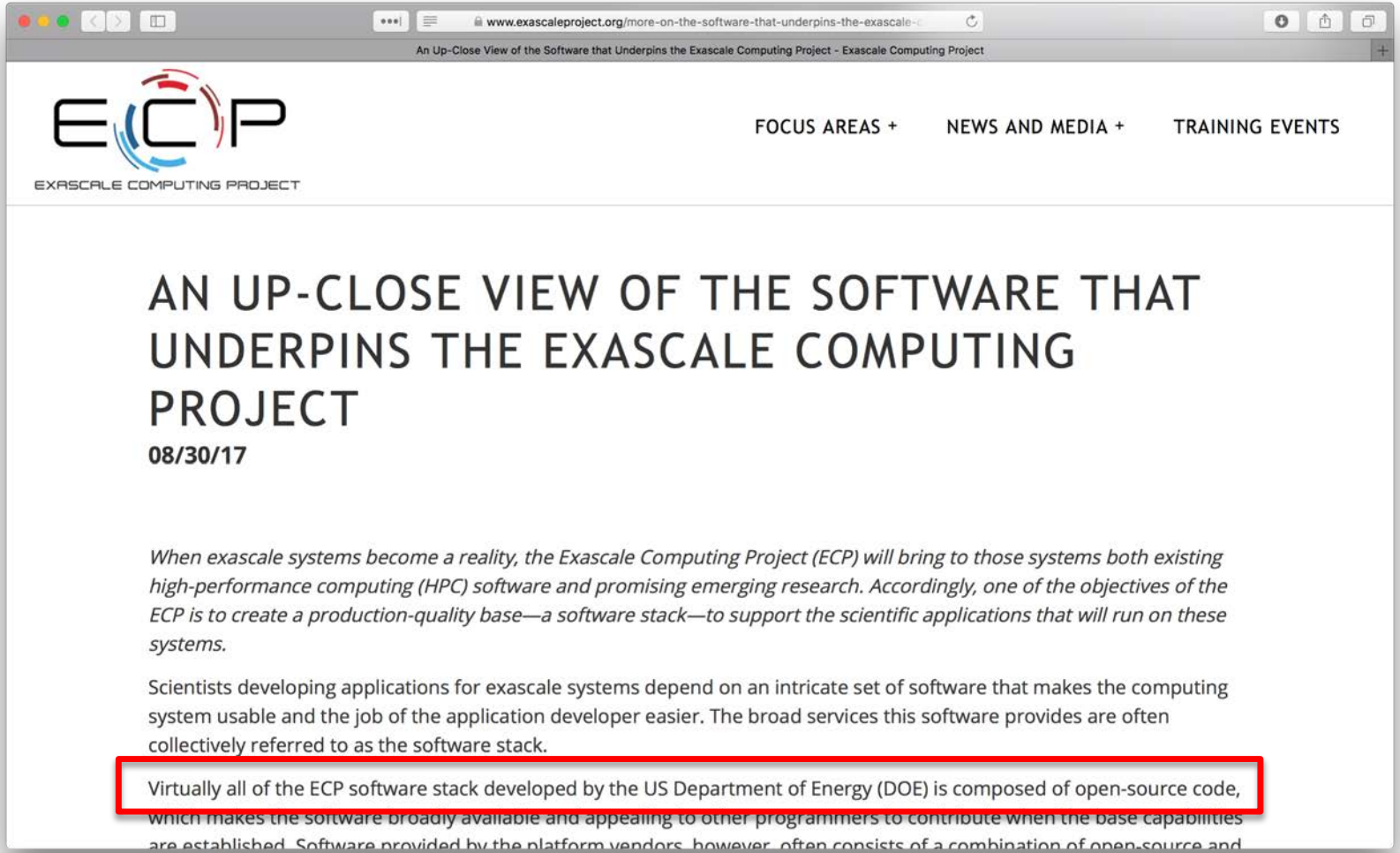



Lawrence Livermore National Laboratory
January/February 2018
Science & Technology
REVIEW
LIVERMORE
SOFTWARE
FOR THE PEOPLE
Also in this issue:
Investigating Bugs in Space
A Code to Simulate Sound
Hydrogen Stored in Solids

Science & Technology Review

"Our large collection of software is a precious Laboratory asset, one that benefits both Lawrence Livermore, and in many cases, the public at large."

- Bruce Hendrickson
Associate Director, Computation


www.exascaleproject.org/more-on-the-software-that-underpins-the-exascale-c
An Up-Close View of the Software that Underpins the Exascale Computing Project - Exascale Computing Project
ECP
EXASCALE COMPUTING PROJECT
FOCUS AREAS +
NEWS AND MEDIA +
TRAINING EVENTS
AN UP-CLOSE VIEW OF THE SOFTWARE THAT UNDERPINS THE EXASCALE COMPUTING PROJECT
08/30/17
When exascale systems become a reality, the Exascale Computing Project (ECP) will bring to those systems both existing high-performance computing (HPC) software and promising emerging research. Accordingly, one of the objectives of the ECP is to create a production-quality base—a software stack—to support the scientific applications that will run on these systems.
Scientists developing applications for exascale systems depend on an intricate set of software that makes the computing system usable and the job of the application developer easier. The broad services this software provides are often collectively referred to as the software stack.
Virtually all of the ECP software stack developed by the US Department of Energy (DOE) is composed of open-source code,

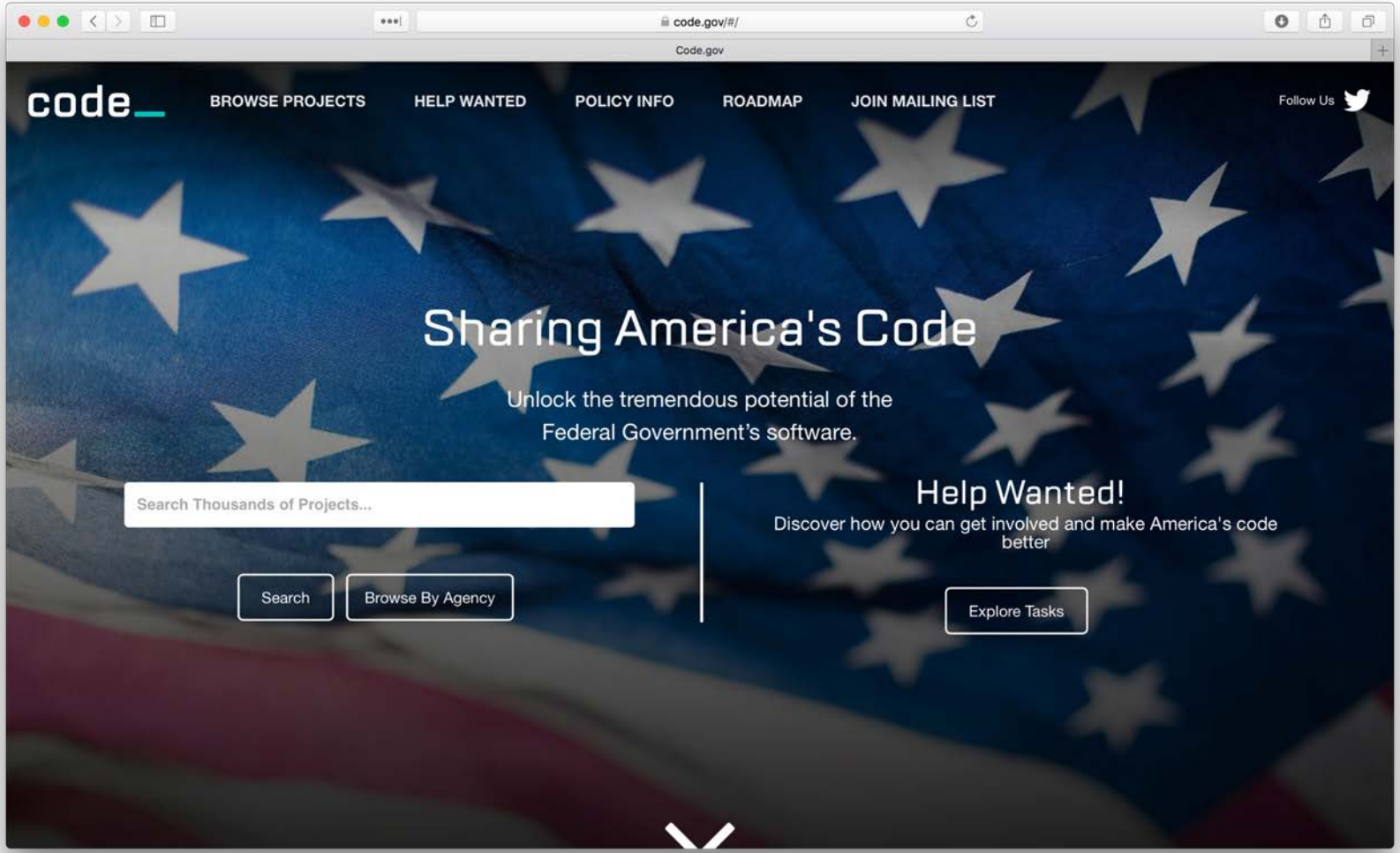
Federal Source Code Policy

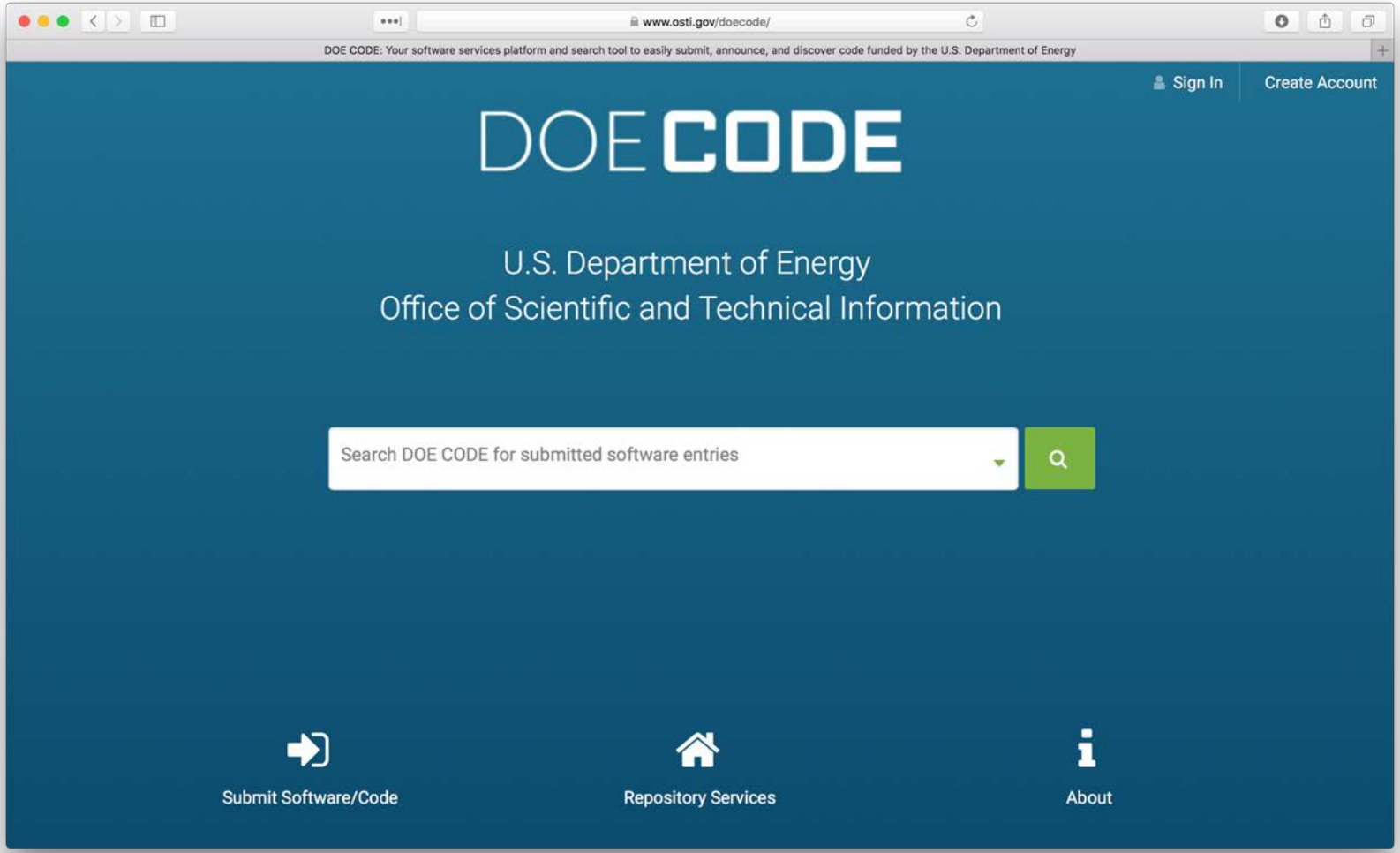
- "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reuseable and Open Source Software"

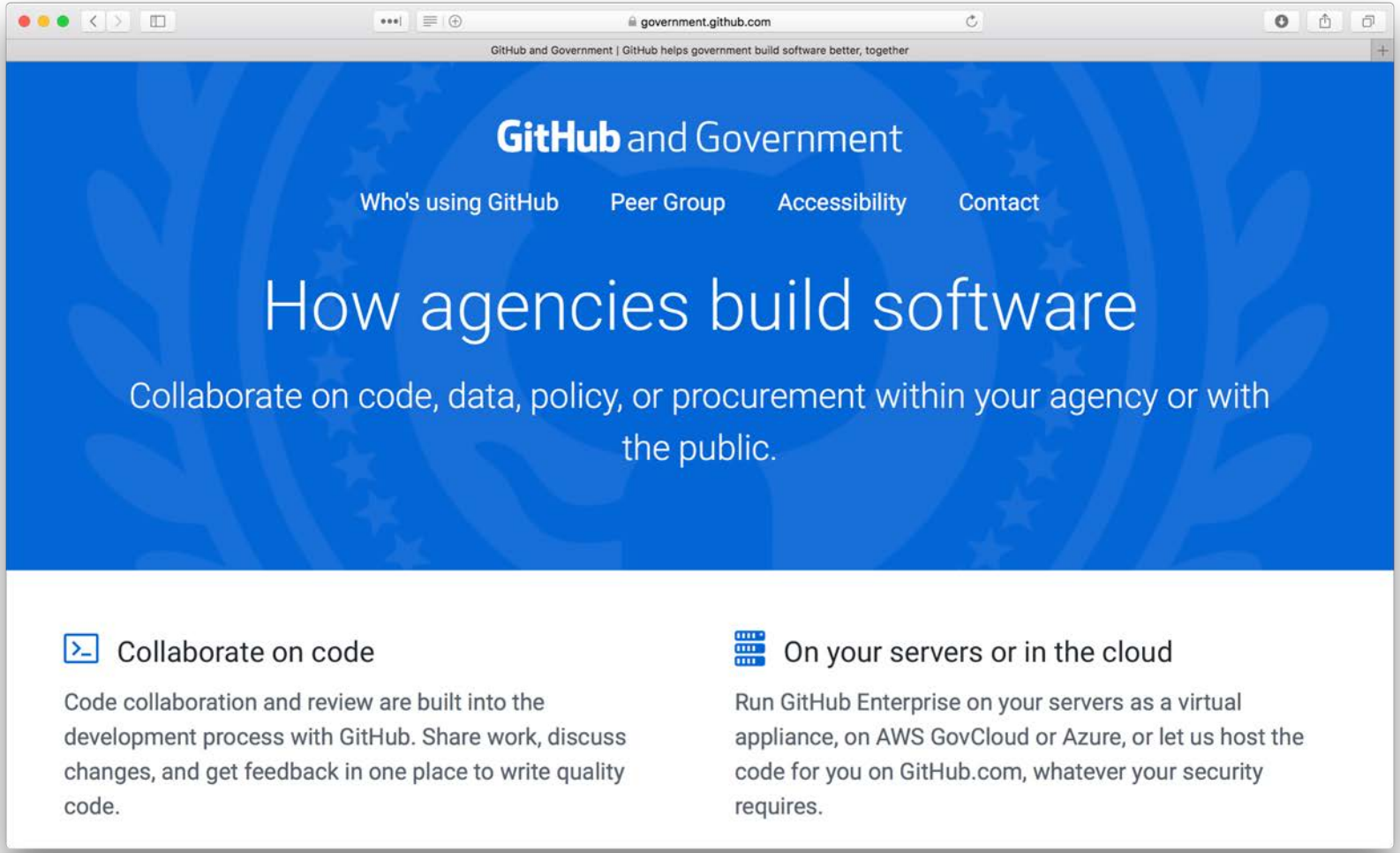
  - "Agencies **shall make custom-developed code available for Government-wide reuse** and make their code inventories discoverable at https://www.code.gov ("Code.gov") […]"

  - "[…] establishes a pilot program that requires agencies, when commissioning new custom software, to **release at least 20 percent of new custom-developed code as Open Source Software** (OSS) […]"

https://code.gov & https://sourcecode.cio.gov


code.gov/#/
Code.gov
code_
BROWSE PROJECTS
HELP WANTED
POLICY INFO
ROADMAP
JOIN MAILING LIST
Follow Us
Sharing America's Code
Unlock the tremendous potential of the Federal Government's software.
Search Thousands of Projects...
Search
Browse By Agency
Help Wanted!
Discover how you can get involved and make America's code better
Explore Tasks


www.osti.gov/doecode/
DOE CODE: Your software services platform and search tool to easily submit, announce, and discover code funded by the U.S. Department of Energy
Sign In
Create Account
DOE CODE
U.S. Department of Energy
Office of Scientific and Technical Information
Search DOE CODE for submitted software entries
Submit Software/Code
Repository Services
About


government.github.com
GitHub and Government | GitHub helps government build software better, together
GitHub and Government
Who's using GitHub
Peer Group
Accessibility
Contact
How agencies build software
Collaborate on code, data, policy, or procurement within your agency or with the public.
Collaborate on code
Code collaboration and review are built into the development process with GitHub. Share work, discuss changes, and get feedback in one place to write quality code.
On your servers or in the cloud
Run GitHub Enterprise on your servers as a virtual appliance, on AWS GovCloud or Azure, or let us host the code for you on GitHub.com, whatever your security requires.

US Government Organizations on GitHub

Thank You!

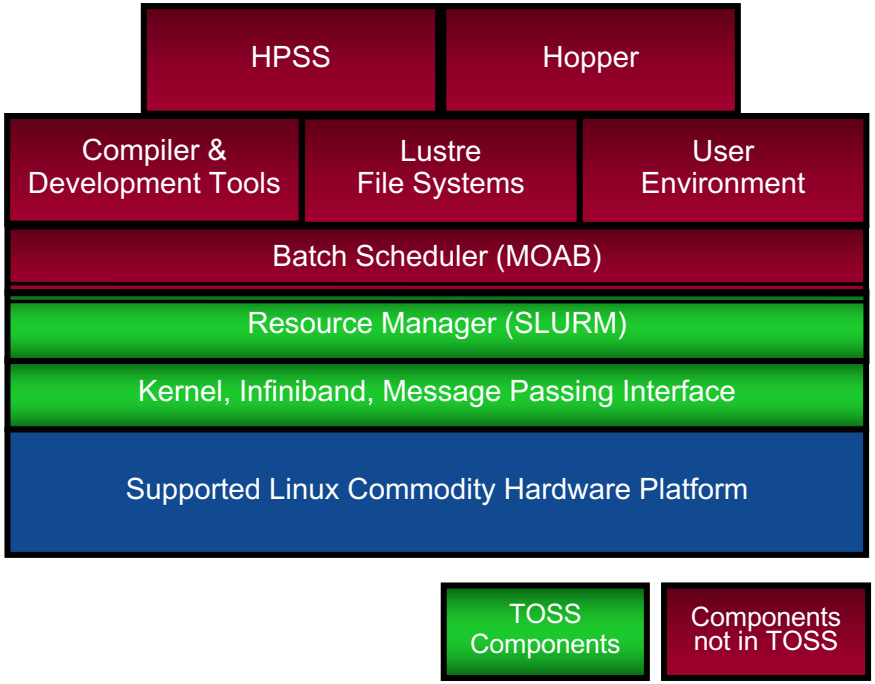
ian@llnl.gov

@IanLee1521 // @LLNL_OpenSource

https://speakerdeck.com/IanLee1521

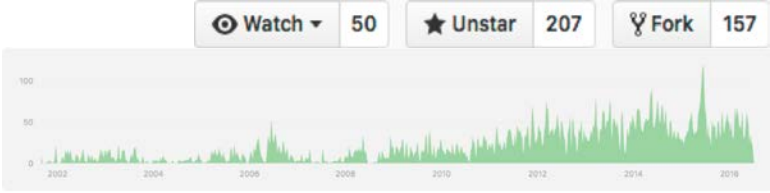
TOSS – Tri-Lab Operating System Software

- Built on Red Hat Enterprise Linux
  - Not an HPC distribution
- Adds LLNL developed additions and patches to support HPC
  - Low Latency Interconnect: Infiniband
  - Parallel File System: Lustre
  - Resource Manager: SLURM
- Work closely with open communities

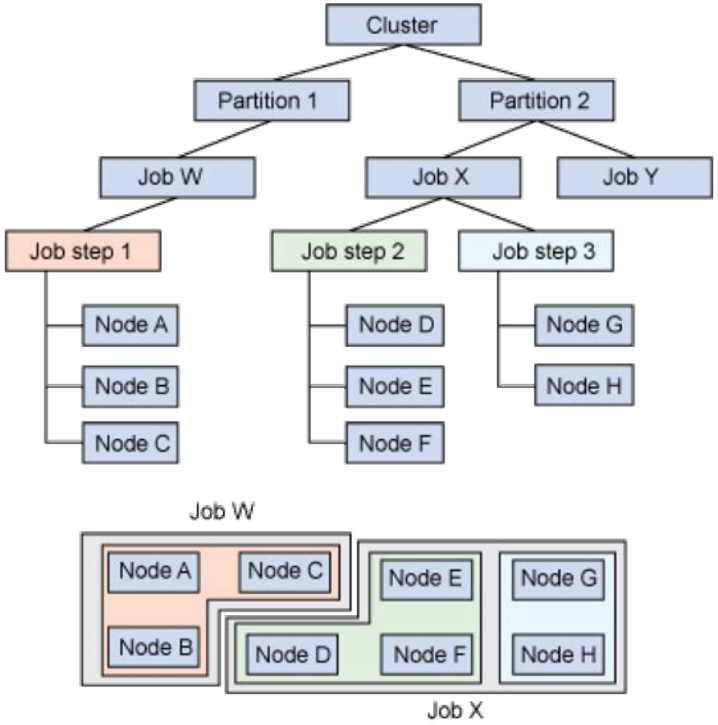


TOSS is a software stack for HPC – large, interconnected clusters!

slurm
workload manager



- Began as simple resource manager
  - Now scalable to 1.6M+ cores (sequoia)
- Launch and manage parallel jobs
  - Large, parallel jobs, often MPI
- Queuing and scheduling of jobs
  - Much more work than resources



http://slurm.schedmd.com

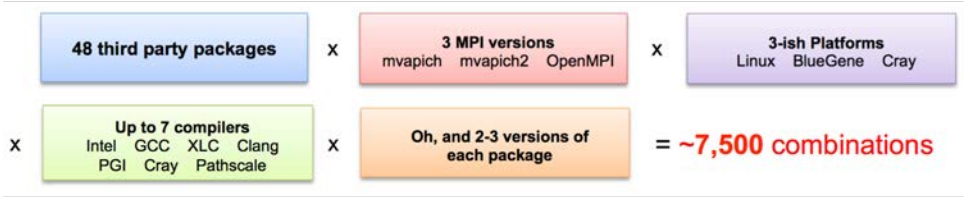
http://slurm.schedmd.com
http://www.ibm.com/developerworks/library/l-slurm-utility/figure3.gif

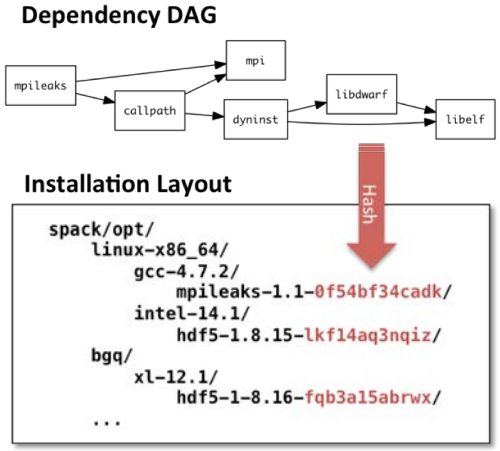
flux

- Family of projects used to build site-customized resource management systems
- flux-core
  - Implements the communication layer and lowest level services and interfaces
- flux-sched
  - Consists of an engine that handles all the functionality common to scheduling
- capacitor
  - A bulk execution manager using flux-core, handles running and monitoring 1000's of jobs

SPACK

- Handles combinatorial explosion of ABI-incompatible packages
- All versions coexist, binaries work *regardless of user's environment*
- Familiar syntax, reminiscent of brew, yum, etc

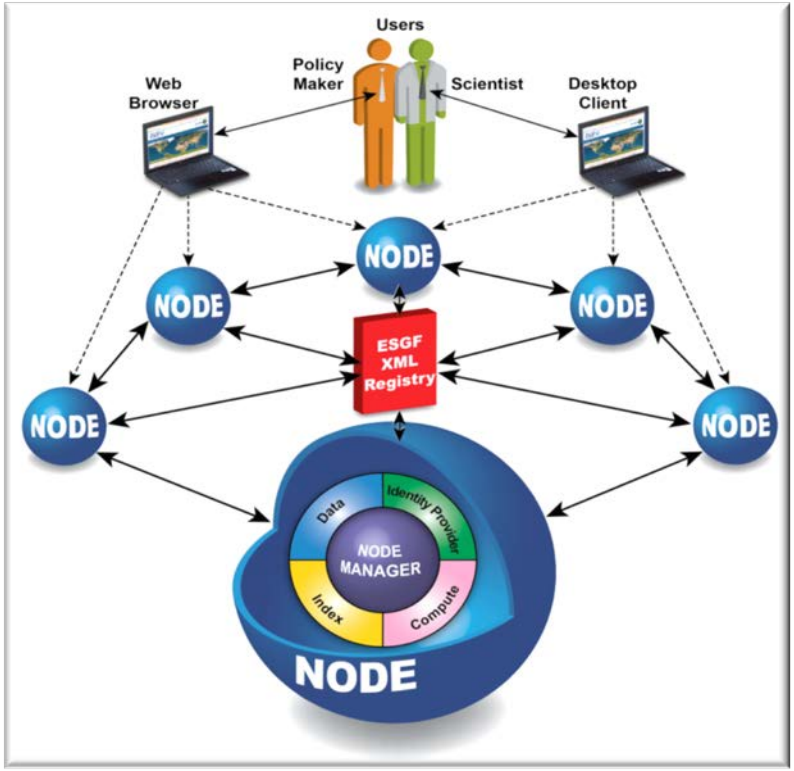





```
$ spack install mpileaks                    unconstrained
$ spack install mpileaks@3.3                @ custom version
$ spack install mpileaks@3.3 %gcc@4.7.3     % custom compiler
$ spack install mpileaks@3.3 %gcc@4.7.3 +threads   +/- build option
$ spack install mpileaks@3.3 os=SuSE11      os=<frontend OS>
$ spack install mpileaks@3.3 os=CNL10       os=<backend OS>
$ spack install mpileaks@3.3 os=CNL10 target=haswell target=<cpu target>
```

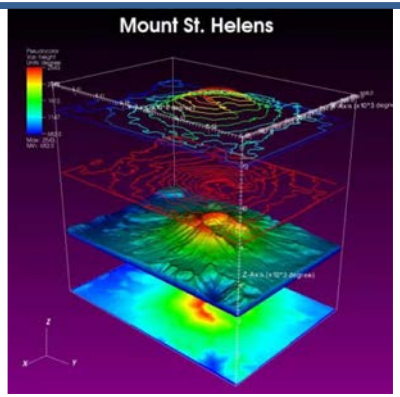
https://github.com/ESGF

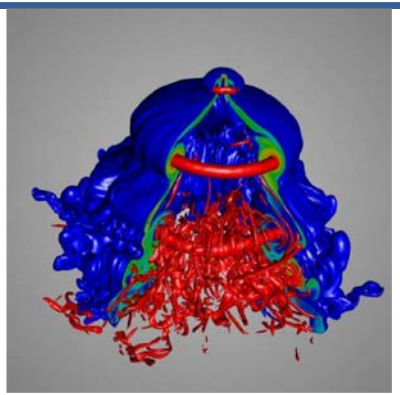
- Manages the first-ever decentralized database for handling climate science data
- Multiple petabytes of data at dozens of federated sites worldwide
- International collaboration for the software that powers most global climate change research

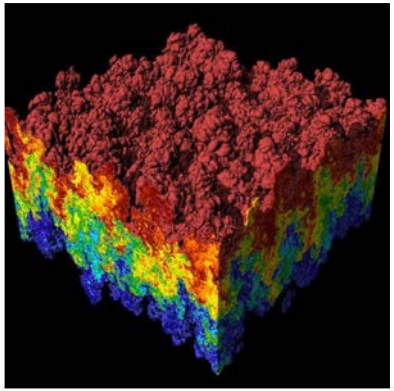


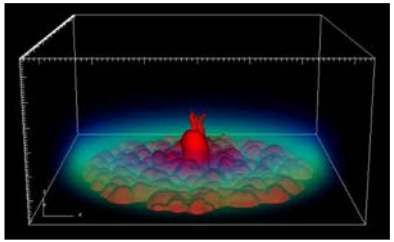
VisIt

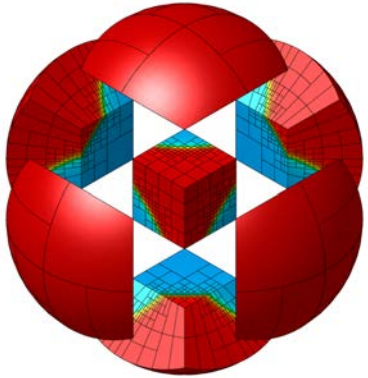
- Originally developed to visualize and analyze the results of terascale simulations
- Interactive, scalable, visualization, animation and analysis tool
- Powerful, easy to use GUI
- Distributed and parallel architecture allows handling extremely large data sets interactively

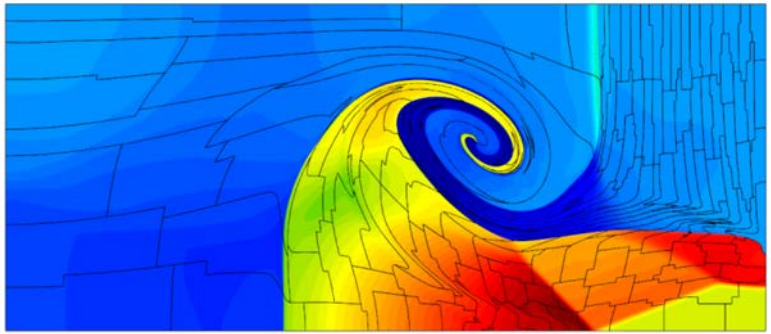


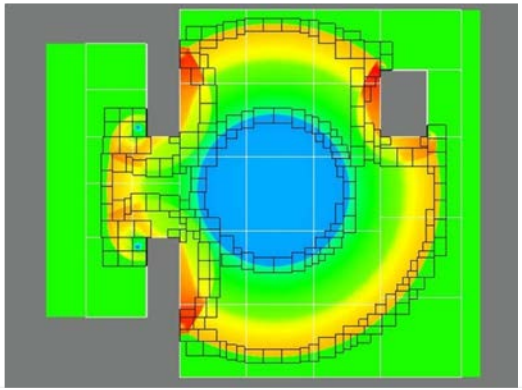


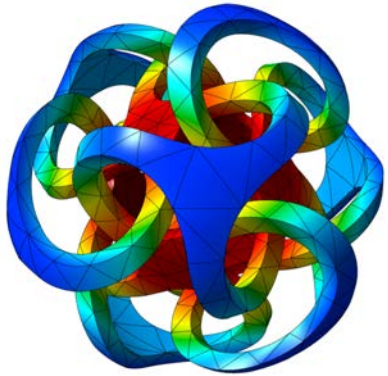


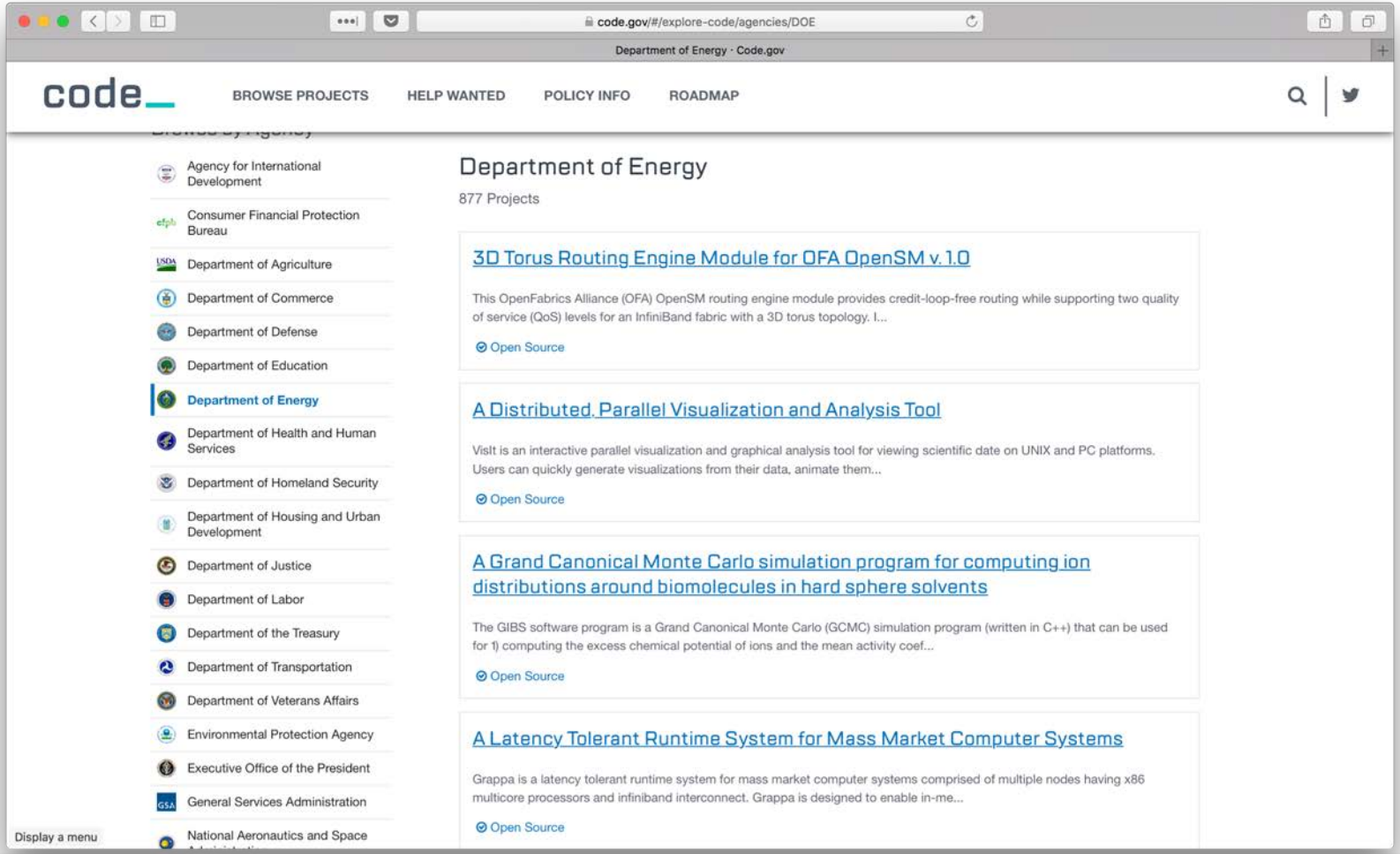











code.gov/#/explore-code/agencies/DOE
Department of Energy · Code.gov
code_
BROWSE PROJECTS
HELP WANTED
POLICY INFO
ROADMAP
Agency for International Development
Consumer Financial Protection Bureau
Department of Agriculture
Department of Commerce
Department of Defense
Department of Education
Department of Energy
Department of Health and Human Services
Department of Homeland Security
Department of Housing and Urban Development
Department of Justice
Department of Labor
Department of the Treasury
Department of Transportation
Department of Veterans Affairs
Environmental Protection Agency
Executive Office of the President
General Services Administration
National Aeronautics and Space
Display a menu
Department of Energy
877 Projects
3D Torus Routing Engine Module for OFA OpenSM v. 1.0
This OpenFabrics Alliance (OFA) OpenSM routing engine module provides credit-loop-free routing while supporting two quality of service (QoS) levels for an InfiniBand fabric with a 3D torus topology. I...
Open Source
A Distributed, Parallel Visualization and Analysis Tool
VisIt is an interactive parallel visualization and graphical analysis tool for viewing scientific date on UNIX and PC platforms. Users can quickly generate visualizations from their data, animate them...
Open Source
A Grand Canonical Monte Carlo simulation program for computing ion distributions around biomolecules in hard sphere solvents
The GIBS software program is a Grand Canonical Monte Carlo (GCMC) simulation program (written in C++) that can be used for 1) computing the excess chemical potential of ions and the mean activity coef...
Open Source
A Latency Tolerant Runtime System for Mass Market Computer Systems
Grappa is a latency tolerant runtime system for mass market computer systems comprised of multiple nodes having x86 multicore processors and infiniband interconnect. Grappa is designed to enable in-me...
Open Source

Public US Government GitHub Data Scrape

- 252 US Government Orgs
  - U.S. Federal (137)
  - U.S. Military and Intelligence (12)
  - U.S. Research Labs (103)

- 8716 Open Source Repositories