University of Colorado Boulder

RMACC - Collaborating Regionally to Facilitate Widespread Effective Use of Advanced Cyberinfrastructure

Thomas Hauser

Director of Research Computing

University of Colorado Boulder

thomas.hauser@colorado.edu

Research Computing
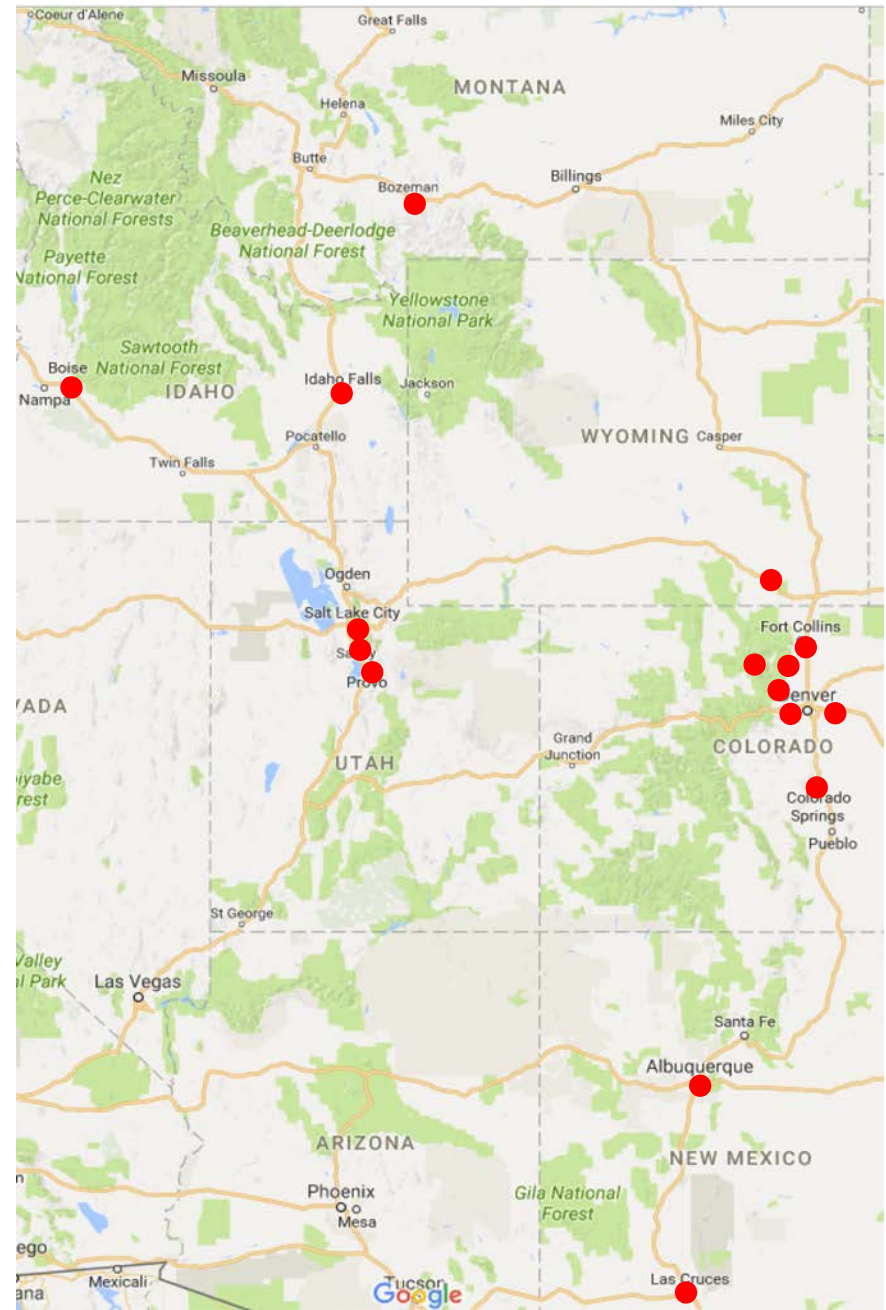UNIVERSITY OF COLORADO **BOULDER**

Be Boulder.

Outline

- RMACC overview
- Community building
- RMACC Summit supercomputer
- Training
- NSF Cyberteam award to support data workflows

Rocky Mountain Advanced Computing Consortium (RMACC)

- www.rmacc.org

- Collaboration among academic and research institutions to facilitate widespread effective use of cyberinfrastructure throughout the Rocky Mountain region

- Membership
  - Full
  - Affiliate

RMACC Leadership

- Chair
  - Thomas Hauser, University of Colorado Boulder

- Vice-Chair
  - Tom Cheatham, University of Utah

- Executive Director
  - Becky Yeager, University of Colorado Boulder

- At Large Members:
  - Tim Kaiser, Colorado School of Mines
  - Diana Dugas, New Mexico State University
  - Ben Nickell, Idaho National Lab

RMACC Activities

- Yearly HPC Symposium
  - Support students with travel grants
  - Poster competition for students (winners will be funded to travel to SC conference)
  - Technical program
  - Tutorials
- Coordinate and collaborate on proposals
- Create a community of system administrators and computational scientists
  - HPC system administrator retreat
  - Working on the facilitators
- Share infrastructure

Research Computing
UNIVERSITY OF COLORADO **BOULDER**

Be Boulder.

RMACC 2018 Symposium

- August 7-9 2018 in Boulder
- Tutorials
  - August 7 & 9
  - Python & R
  - HPC & cloud computing
- Technical Program
  - August 8
  - Student poster competition
  - Data management
  - Security
  - Big data & machine learning

Research Computing
UNIVERSITY OF COLORADO **BOULDER**

Be Boulder.

Outreach

- Outreach committee
- First pilot in New Mexico
  - Event at a smaller school
  - Meeting with CIO and other campus leadership
  - Meeting with faculty and students to learn about their needs
  - New RMACC Summit user seminar
- Goal is to facilitate 2-3 events a year
- Working on defining metrics of success

System Administrator Community

- 3 times a year in person meeting
  - RMACC symposium
  - SC conference series
  - Additional meeting at one of the member institutions
- April meeting at the University of New Mexico
  - Private cloud
  - Containers
  - Students
  - Cooling

Summit: RMACC's Supercomputer

- Funded via an NSF MRI grant awarded jointly to CU Boulder and CSU
- Available to any RMACC member researcher
  - 10% to RMACC
  - Rest to CSU and CSU proportional to the award amount
- Theoretical peak performance around 450 TFLOPS
- Mix of established and cutting-edge technology
- 5-year lifetime
  - Just started year 2
- Principal vendor and integrator is Dell

Research Computing
UNIVERSITY OF COLORADO **BOULDER**

Be Boulder.

RMACC Summit

- 380 general compute nodes

- 10 GPGPU/visualization nodes
  - 2x NVIDIA K80 GPUs

- 5 High-memory nodes
  - 2 TB RAM each

- 20 Xeon Phi ("Knights Landing") nodes
  - Phi installed directly in CPU socket; not a separate card
  - 72 cores / 288 threads   "many core"

- GPFS for scratch file system

- Omnipath as the HPC interconnect.

Training

- RC at CU Boulder offers trainings virtually or in person

- Home grown or from external sources

- Open to the general public
  - One of our missions is to educate our community

- Measuring training impact
  - Survey
  - User data

- Regional training offerings
  - Utah

NSF CC*Cyber Team: Creating a Community of Regional Data and Workflow Cyberinfrastructure Facilitators

PI: Thomas Hauser, CU Boulder

Co-PIs: Pat Burns, H.J. Siegel, CSU

Tom Cheatham, University of Utah

Research Computing
UNIVERSITY OF COLORADO **BOULDER**

Be Boulder.

Focus Areas

- Data curation, preservation and metadata
  - CSU

- Protected data and compute workflows
  - University of Utah

- Data and compute workflows
  - CU Boulder
  - Integrated in the Center for Research Data and Digital Scholarship

Approach

- Create standards for engagement with science teams
    - Based on ESnet approach
    - Follow-up to assess impact
    - IRB approval
- Community building
- Professional development for facilitators
- Training for researchers
- Regional shared cyberinfrastructure
    - Shared data catalog

Impacts

- Focus on end to end user workflows
  - User analytics using data
    - Bash history
    - Auditd data
    - Networking data

- Institutionalize culture of data management

- Support compliance with federal mandate for data management plans

- Share expertise, strategies and techniques with RMACC partners

- Support smaller institutions in the RMACC regions

Questions?

University of Colorado **Boulder**