Computational Science Research and Education Initiatives at the UNM

Patrick G. Bridges
Director, UNM Center for Advanced Research Computing
Professor, Department of Computer Science
Mid-size Academic HPC

- How to create a sustainable research computing ecosystem that serves the state’s research, education, and workforce development needs?

- Research
  - Lots of “long-tail” small-scale research in non-traditional data and computing areas
  - Mixed with sophisticated users develop towards national facilities

- Education
  - Research computing increasingly prevalent in disciplines without a history to build on, teach from
  - Modern economy demands a broadly data- and compute-literate public!

- Workforce development: Need coherent pathways from K-12 through graduate to meet state- and nation-wide needs
Mid-size Academic HPC Challenges

- **Traditional systems ill-suited to workloads**
  - How to support non-traditional users and users working towards national facilities with diverse environments?
  - How to handle data availability, preservation requirements?
  - How to handle CUI data and compute without over-burdening support staff, users?

- **Academic computing budgets scarce**
  - Lots of small-scale investments (tens of thousands) spread out haphazardly over time
  - Occasional moderate infusions (hundreds of thousands+) from institution, partners, grants

- **Support and background generally poor**
  - Researchers, lab, and departmental staff have little to no experience with research computing state of the art best practices
  - Attracting and retaining qualified staff incredibly difficult
UNM Computational Science Research and Education Initiatives

- **Computational Science Research**
  - Cutting edge research by multiple faculty in diverse departments
  - Areas range from traditional (e.g. CFD) to emerging (quantum computing) to non-traditional (e.g. forensic anthropology)

- **Cyberinfrastructure Research**
  - Designing new systems for effective, sustainable academic research computing

- **UNM Education and Workforce Development**
  - Huge need for HPC and computational science expertise nation-wide and in New Mexico
  - Working to integrate and expand a wide range of programs ranging from K-12 through graduate
Computational and Data-Intensive Science Research Examples

- Mountain Lions on the Edge: Integrating Conservation into Urban Planning through Predictive Modeling – Prof. Bruce Milne, UNM Department of Biology
- Performance Optimization of LANL Multi-Physics Applications – Prof. Patrick Bridges, UNM Department of Computer Science
- A High-fidelity Model for Wind Farms – Prof. Sang Lee, UNM Department of Mechanical Engineering
- Economic Network Simulations – Prof. David Dixon, UNM Department of Economics
- New Mexico Decedent Information Database – Prof. Heather Edgar, UNM Anthropology
- Multiscale Mechanistic Model to Study Nanotherapy Delivery in Tumors – Dr. Elaine Bearer, UNM HSC Department of Pathology
Computational and Data-Intensive Science Research Examples

- Mountain Lions on the Edge: Integrating Conservation into Urban Planning through Predictive Modeling – Prof. Bruce Milne, UNM Department of Biology

- Performance Optimization of LANL Multi-Physics Applications – Prof. Patrick Bridges, UNM Department of Computer Science

- A High-fidelity Model for Wind Farms – Prof. Sang Lee, UNM Department of Mechanical Engineering

- Economic Network Simulations – Prof. David Dixon, UNM Department of Economics

- New Mexico Decedent Information Database – Prof. Heather Edgar, UNM Anthropology

- Multiscale Mechanistic Model to Study Nanotherapy Delivery in Tumors – Dr. Elaine Bearer, UNM HSC Department of Pathology
Phantom-Cell AMR

- Working with LANL on a new hybrid approach to adaptive mesh refinement that:
  - Allows users to run regular physics codes on irregular meshes
  - Allow simple portable code without refinement or GPU portability problems of patch/block-based AMR

- Working demonstration in LANL CLAMR mini-application
  - Faster runtimes, simpler code, potential easier porting to GPUs
  - Increased memory usage (due to data structures to handle mesh mapping)
NM Decedent Information Database

- **Database of decedent information**
  - Primary data from the NM Office of the Medical Investigator
  - Augmented with next-of-kin survey data, medical information summaries
  - 15,000 full-body CT scans (150,000,000+ images)

- **Research uses in forensics, medical image analysis, anthropology, etc.**

- **Development and data transfer in progress**

- **Funding from US National Institutes of Justice**

- **Exploring many possible next research and dataset steps**
Systems supporting this research

■ Range of local systems
  ▪ 300-node/2400-core capacity compute cluster (donated from LANL via NSF PROBE program)
  ▪ 32-node GPU cluster for imaging, machine learning, big memory workloads (dual GPUs, 2 each 1TB and 3TB RAM nodes)
  ▪ Condo-model cluster for users with specialized computing needs
  ▪ Storage and VM-hosting systems (primarily NetApp and VMWare)

■ UNM users also transition to a range of national resources
  ▪ NSF XSEDE systems (primarily TACC systems)
  ▪ NERSC, NREL, DoD systems
  ▪ Some work with national lab collaborators to execute on more specialized, limited access systems
  ▪ National lab sponsors often have students use UNM systems for development due to short queues
How do we design cyberinfrastructure systems to support this environment?

- Broadly embracing system virtualization and containerization
  - Improves and eases incremental system growth and management
  - Compute and storage re-provisioning with long-term workload changes
  - Can we import external compute environments into this space to ease portability, repeatability?
  - NSF CiCi: Create custom virtually isolated environments for CUI needs
  - Dell/VMWare/Mellanox: How to effectively virtualize parallel storage systems?

- Beginning to integrate with institutional enterprise IT units
  - System needs are increasingly similar
  - Challenges between stability/compliance-driven enterprise computing and agility-driven research computing cultures and processes
New Mexico Data to Insight

Vision: New Mexico needs a data-literate populace that can effectively use next-generation data and compute tools and techniques to make insightful, informed day-to-day decisions.

- **Goal:** Create educational pathways for the people of New Mexico for long-term workforce and economic development

- **Approach:** Integrate diverse UNM computing and data research and education programs with new initiatives into a coherent program to teach people to turn data into day-to-day insights.
  - **K-12 Programs:** Basic computational and data literacy in the schools
  - **General undergraduate and graduate:** Integrate education in computational thinking across a range of disciplines
  - **Specialized undergraduate and graduate:** Educate the next generation of compute and data scientists, mentor research computing best practices
  - **Professional development:** train existing teachers and professionals in state-of-the-art computing education and research practices
K-12 Education Programs

- **Bosque Ecosystem Monitoring Program**
  - K-12 and university students and school teachers in the collection and analysis of local ecological data
  - [www.bemp.org](http://www.bemp.org)

- **Advancing Out-of-School Learning in Mathematics and Engineering (AOLME)**
  - An integrated mathematics and computer programming curriculum for urban and rural bilingual middle schools
  - [https://aolme.unm.edu/](https://aolme.unm.edu/)

- **New Mexico CS4All**
  - On-campus and distance education in CS education and basic computational modeling for high school teachers
  - [https://cs4all.cs.unm.edu/](https://cs4all.cs.unm.edu/)

- **New Mexico Supercomputing Challenge**
General Undergraduate and Graduate Programs

- Challenge: integrate computing and data science into other disciplines
- Developing courses, course modules, and computational and data science projects for use by students in these disciplines
- Collaboration with consortia of NM colleges and universities
- DOD support for UNM Critical Technologies Studies Program
  - 2 day symposium at UNM next week with national and UN speakers
  - [https://ctsp.unm.edu/annual-symposium/index.html](https://ctsp.unm.edu/annual-symposium/index.html)
- Proposals to NNSA, NSF to fund additional course and project development, student stipends, expansion to additional NM partner institutions (NMT, NMSU, SIPI, others)
Integrating Specialty HPC/Data Science Knowledge into Graduate Research Education

- **Challenge:** Teaching discipline graduate students best practices in HPC and data sciences

- **Approach:** Providing half-time internship (10 hours/wk) at UNM CARC to RAs in other departments
  - Early-program students work hand-in-hand with HPC center staff to learn how to effectively use HPC systems to support their research
  - Help develop discipline-specific documentation, tools, workflows, course modules to support researchers in their lab and field
  - Return full time to their research lab at internship completion, spreading specialized HPC expertise across campus

- **Looking to expand to include data management trainings in collaboration with UNM Libraries and Learning Sciences**
Specialized Workforce/ Professional Development Programs

- **Challenge:** Professional development/retraining for National Lab staff

- **Revamping UNM’s Computational Science and Engineering (CSE) graduate certificate program**
  - Augment existing or ongoing graduate education with additional expertise in computational science and engineering
  - 4 graduate courses plus a compute/data-intensive research project

- **Program evolving to sit at the intersection of computational sciences and data sciences**
  - New programmatic changes to make data science a core portion of the curriculum across all computational science areas
  - Creating specializations in multiple areas of local and national need
    - Data Science
    - Computational Fluid Dynamics
    - Quantum Information Systems

- **Looking at developing an online version of the program**
Broader Education Challenges

Integrating these varied programs into a coherent set of pathways for diverse student populations is challenging!

- Identify and bridge gaps between programs
  - Many potential students lost K-12 to undergraduate transition
  - Internships can be a good mechanism for bridging these gaps

- Need appropriate peer-to-peer and near-peer mentoring
  - Undergrads mentor high school students, graduate students under graduates, and lab personnel graduate students?

- Provide multiple well-defined pathways through existing programs for prospective students
Questions?

Prof. Patrick Bridges
University of New Mexico
patrickb@unm.edu