INFRASTRUCTURE FOR HIGH PERFORMANCE COMPUTING

Amy Apon, Ph.D.
Director, Arkansas High Performance Computing Center
Professor, Computer Science and Computer Engineering
University of Arkansas

Work supported in part by NSF Grant #0918070
and by Governor Beebe through the Arkansas
Science and Technology Authority

CYBERINFRASTRUCTURE

"The IT infrastructure that enables scientific inquiry" – Daniel Atkins

Research
Nanotechnology
Geoinformatics
and more…

Education
Computational science, engineering, math, and technology

Telemedicine
Arkansas Teleheath Network

Arkansas Cyber-infrastructure
- ARE-ON
- High Performance Computing
- TEOS

Service Providers

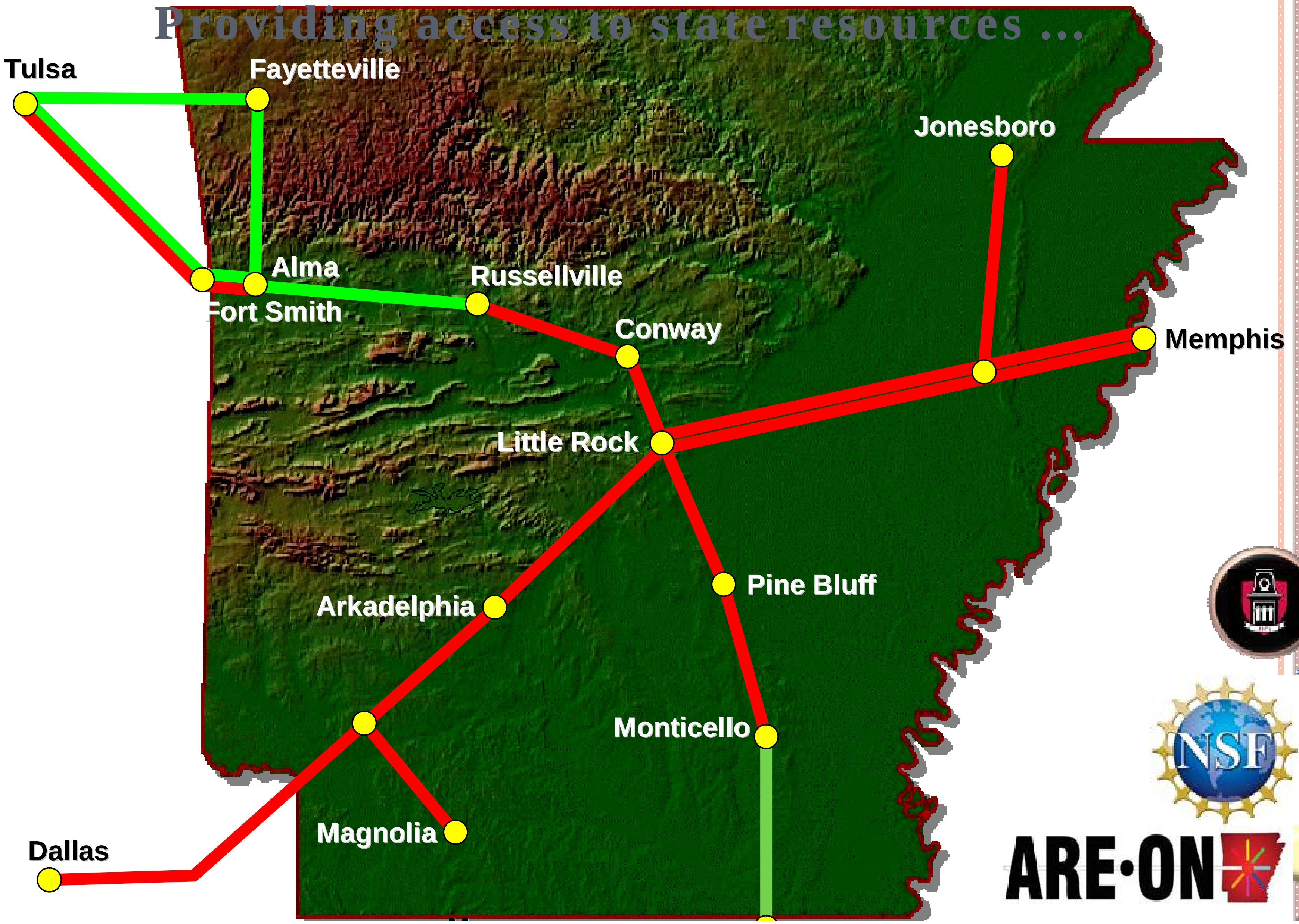
Emergency Preparedness

Business
Precision farming
Large scale business modeling
and more…

Connect Arkansas
Planning & Outreach

ARKANSAS RESEARCH AND EDUCATION OPTICAL NETWORK

Providing access to state resources …



… and access to national cyberinfrastructure resources.

† NSF Office of Cyberinfrastructure, Rob Pennington

SUPERCOMPUTER

Funded through NSF MRI #0722625 #339, June 2008, 1256 cores, 10.75Tflop/s, 11 million compute hours/year

Other current state resources combined total ~500 cores.

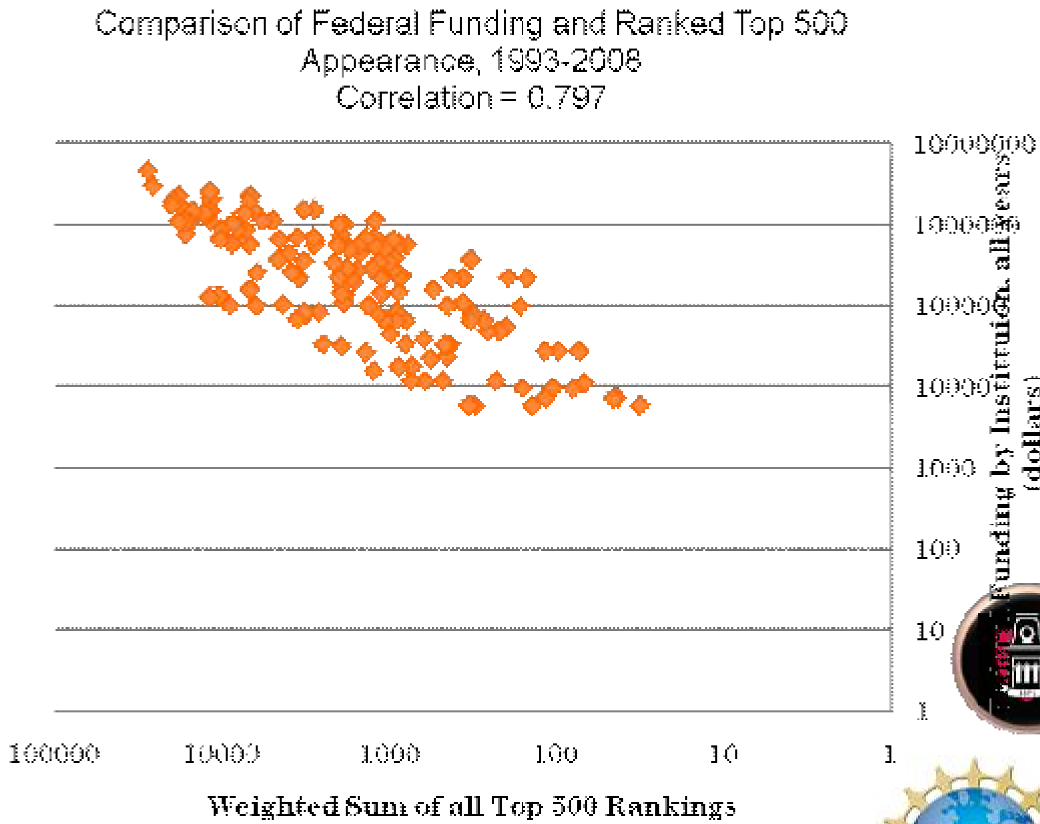
MAJOR RESEARCH AREAS

- Current computational research in Arkansas includes
  - Computer science
  - Chemistry
  - Physics
  - Materials science
  - Electrical engineering
  - Geosciences

COMPLEX DATA ANALYSIS USING EMERGING TECHNOLOGIES

<u>The challenge</u>

- There is a lot of data, from diverse sources! New multicore and accelerator architectures provide new opportunities for fast data processing.
- Research by Amy Apon, and colleagues develop programming models on emerging technologies for advanced large scale data processing.

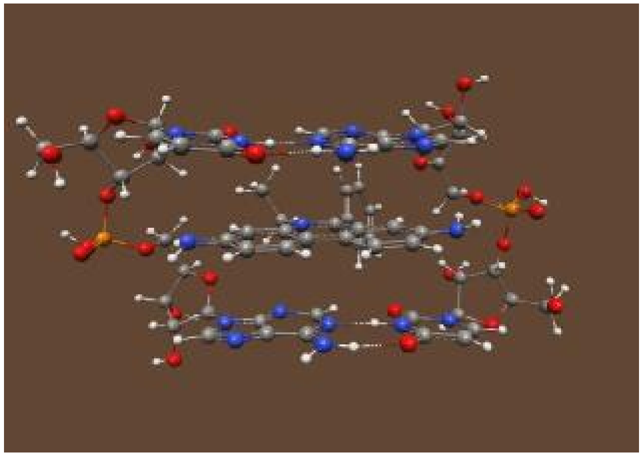


ACCURATE CALCULATION OF LARGE MOLECULES

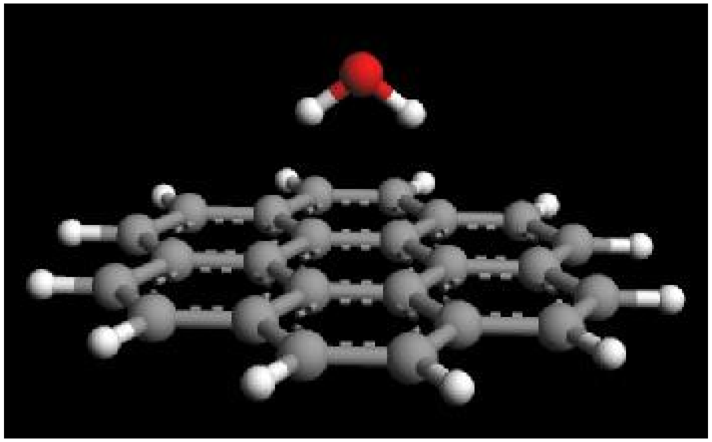
The challenge

We want to understand cancerogenic processes that occur in our environment.

- Research by Peter Pulay and his team study the interaction of chemicals on human protein, and DNA structure
- Pulay's research requires 4 million hours of compute time each year



Ethidium Bormide between two DNA strands



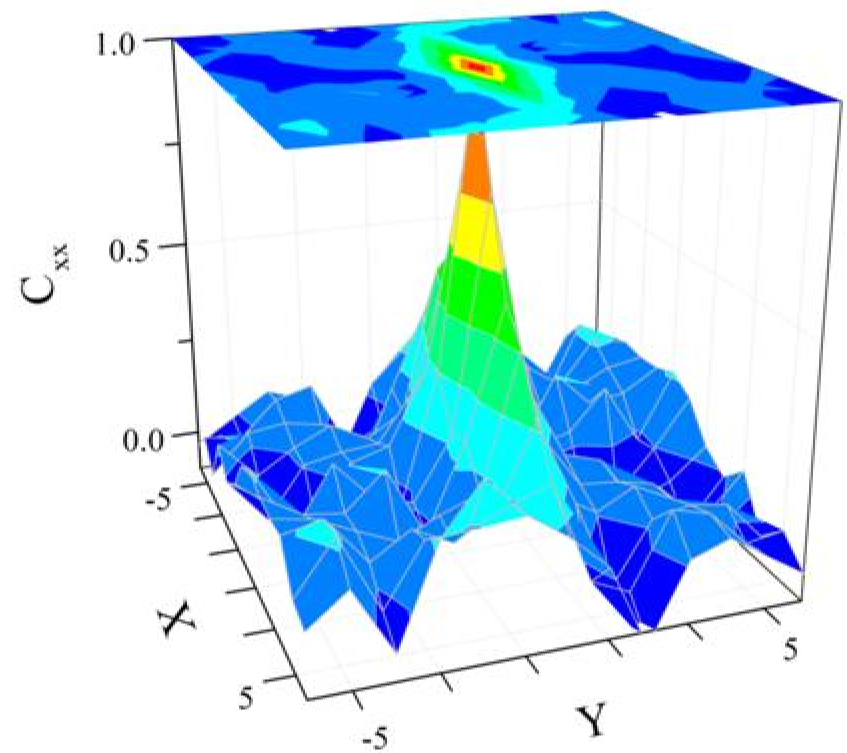
Water and a graphitic surface

COMPUTATIONAL NANOTECHNOLOGY

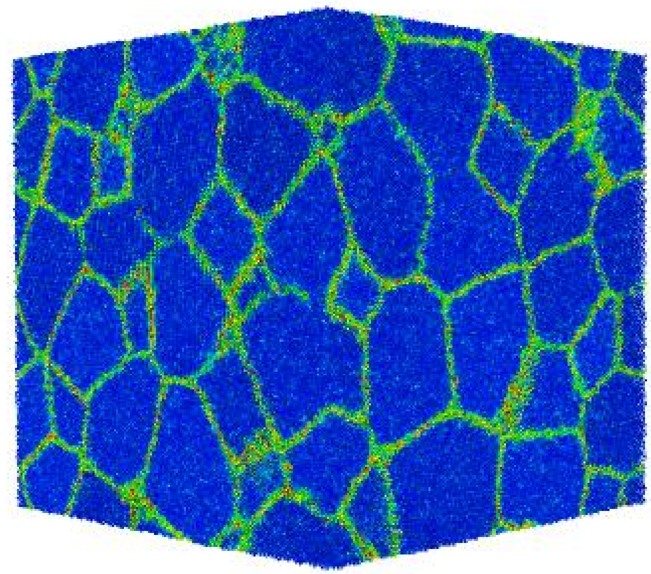
<u>The challenge</u>

We have reached fundamental limits in computer technology.

- Research by Laurent Bellaiche creates nanotechnology devices that can build memory10,000 time denser than what is currently manufactured.
- Bellaiche's research requires 70 million hours of compute time each year.



MATERIALS SCIENCE



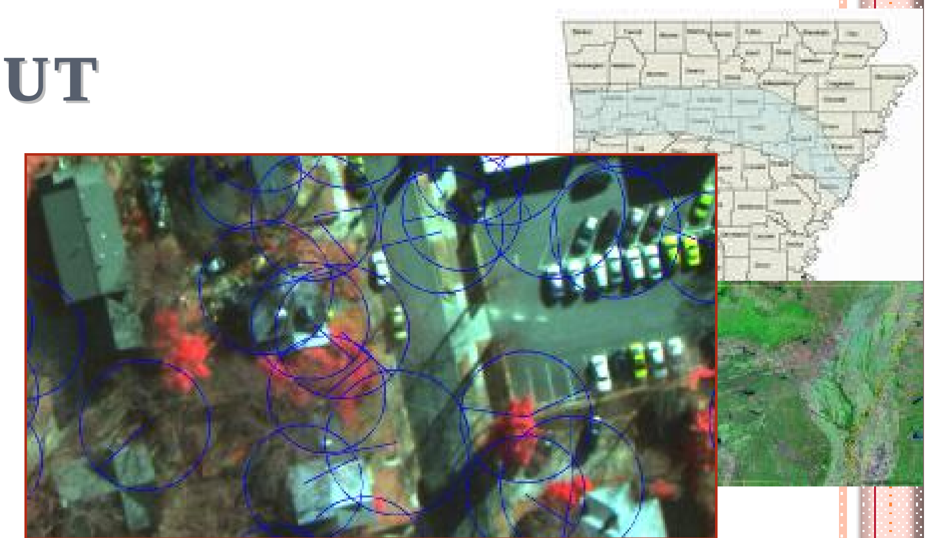
The challenge: model plasticity and failure in metal alloys – this has applications in aeronautics

- Doug Spearot, Assistant Professor of Mechanical Engineering, creates 3-dimensional models using 20 million (or more) atoms.
- These calculations evaluate alloys with different compositions BEFORE they are fabricated in a laboratory
- Computational materials science research requires 6 million hours of compute time each year.

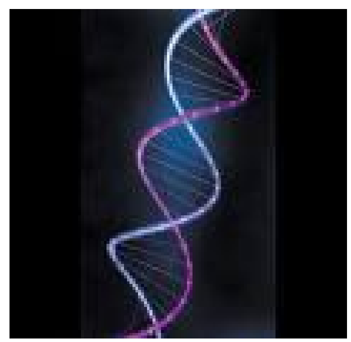
ARKANSAS HAS MANY AREAS OF RESEARCH THAT CANNOT BE DONE WITHOUT SUPERCOMPUTING

"Over the past 60 years, computing has become the most important general-purpose instrument of science"

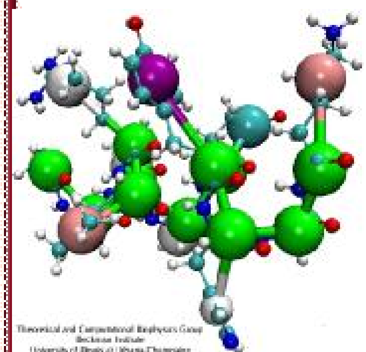
– Jay Boisseau, Director, Texas Advanced Computing Center



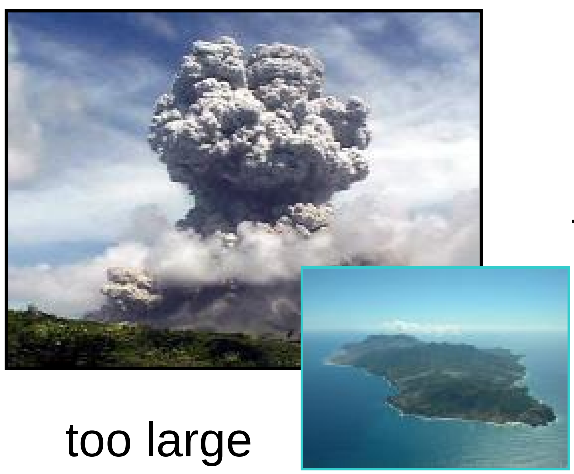
too much data



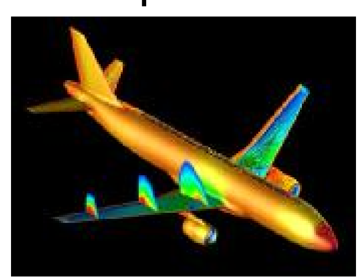
too complex



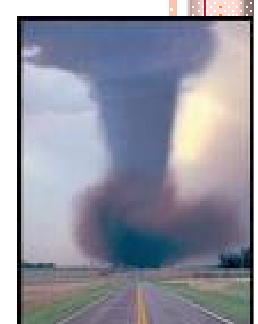
too small



too large



too dangerous to manipulate



SUPERCOMPUTING USES PARALLELISM

Using these tools effectively is difficult

CI-TRAIN

Cyberinfrastructure for Transformational Scientific Discovery

A new funded project from NSF

- Faculty and staff *Campus Cyberinfrastructure Champions* serve as **liaisons** between researchers and resources, both local and national,
- using shared *large-scale computational and visualization resources* and high-speed network access to participating institutions
- to SUPPORT RESEARCH in a wide spectrum of computational and visualization domains.

http://www.ci-train.org/

PROJECT COMPONENTS

- Grow the cyber-workforce
  - Initiatives at the high school, college undergraduate, and graduate levels, with professional information technology staff, and research faculty.
  - Partners at UALR (Srini Ramaswamy), UAPB, UAMS, ASU, and others
- Shared nationally competitive visualization resources
  - will enable a broad range of scientific research and educational activities across several computational science and engineering domains
- Shared large-scale computational resources

KEY CI TRAIN ACTIVITIES IN YEAR ONE

- SC09 Workshop, "Introduction to Computational Thinking," UAF campus, August 2009
- CI Training Days weekly teleconferences – includes CI Campus Champions from 8 campuses
- "Introduction to Scientific Visualization," TACC, January 2010, attended by 23 project members
- CI Days, UAF campus, May 16-17, 2010
  - CI Days events being planned for all participating campuses in succeeding years
- Virtual Summer School – HD stream to UAF campus (Petascale, Large Data, CUDA), summer 2010
- CI TRAIN Yearly Project meeting, in conjunction with TeraGrid '10 at PSC, August, 2010

CRITICAL OPPORTUNITIES AND ISSUES

- HPC tools and predictive simulations can potentially transform science and engineering
- Multi and many core CPUs and hierarchical memory structures have dramatically changed computing
- New grants from NSF (CI TRAIN, MRI) will provide funds for ~2000 new cores, storage, and visualization
- Usability of these resources will need
  - Program development environments, libraries, novel algorithm advances, application software, and collaboration tools and environments
  - **Unprecedented levels of sophistication for computer, data, visualization**
  - We need to seriously rethink our campus environments and how they can support new data-driven modalities of research, collaboration, and education †

† NSF Office of Cyberinfrastructure, Rob Pennington

NEEDS FROM AN OCI PERSPECTIVE †

- **Must educate students at all levels in collaborative computational science**
  - Example problem: Grad students do not understand software & HPC, nor do their advisors
  - Example solution: send them to labs/centers for summer
- **Must encourage/support researchers to move into these areas**
  - Example problem: Recent computer science PhD disconnected from what scientists need, recent physics PhD not trained in software engineering
  - Example solution: create postdoc-to-professoriate programs to encourage them to apply their knowledge (& protect them)
- **Must catalyze culture changes in academia and agencies to better support these activities**
  - Example problem: Computational scientists considered peripheral in many traditional departments, even though they are central to future
  - Example solutions: Consider joint appointments specifically for computational scientists

† NSF Office of Cyberinfrastructure, Rob Pennington

Amy Apon, Ph.D.

aapon@uark.edu

Arkansas High Performance Computing Center (AHPCC)

http://hpc.uark.edu

CI-TRAIN

Cyberinfrastructure for Transformational Scientific Discovery

http://www.ci-train.org/