A Case Study for HPC Workforce Development and Workforce Meta-Development
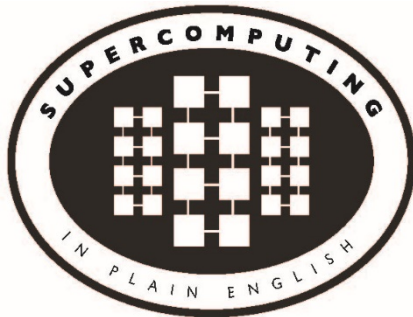
Henry Neeman
Director, OU Supercomputing Center for Education & Research (OSCER)
Assistant Vice President, Information Technology – Research Strategy Advisor
Associate Professor, College of Engineering
Adjunct Associate Professor, School of Computer Science
University of Oklahoma
HP-CAST 2015, Fri Nov 13 2015

HPC Workforce Development is Hard

Why is HPC Workforce Dev Hard?

- At the national scale, no academic discipline wants to teach HPC:
  - Obvious home: Computer Science.
  - Most Computer Science faculty aren't interested in HPC.
    - From their perspective, research about HPC is a tapped well.
    - CS cares about big-O, but HPC is about the constant.
  - Most CS departments don't want to teach service courses primarily for the benefit of other disciplines.

Why is CSE Workforce Dev Hard?

- At the national scale, no academic discipline wants to teach Computational Science & Engineering (CSE).
    - At most institutions, if you do CSE in a CS department, your CS colleagues don't know how to judge your papers that are published in non-CS venues.
    - At most institutions, if you do CSE in a STEM domain department (e.g., Chemistry, Industrial Engineering), your department colleagues don't know how to judge your papers that are published in journals outside your STEM domain discipline.
    - So, at most institutions, you probably won't get tenure.
    - In which case, you won't develop the current workforce, let alone the next generation of workforce developers (faculty).

OneOklahoma Cyberinfrastructure Initiative (OneOCII)

OneOK Cyberinfrastructure Initiative

- All academic institutions in Oklahoma are eligible to sign up for free use of OU's and OSU's centrally-owned Cyberinfrastructure resources.

- Other kinds of institutions (government, non-governmental) are eligible to use, though not necessarily for free.

- Everyone can participate in our CI education initiative.

- The Oklahoma Supercomputing Symposium, our annual conference, continues to be offered to all.

- Triggered by OK's NSF EPSCoR Research Infrastructure Improvement Track-1 2008-13.

OneOCII Goals

- **<u>Reach</u>** institutions outside the mainstream of advanced computing.

- **<u>Serve</u>** every higher education institution in Oklahoma that has relevant curricula.

- **<u>Educate</u>** Oklahomans about advanced computing.

- **<u>Attract</u>** underrepresented populations and institution types into advanced computing.

OneOCII Service Methodologies Part 1

- **<u>Access (A)</u>**: to supercomputers and related technologies (20 OK academic institutions to date).

- **<u>Dissemination (D)</u>**: Oklahoma Supercomputing Symposium – annual advanced computing conference has reached 112 academic institutions, 143 commercial, 36 government, 20 nongovernmental (25 OK academic institutions to date).

- **<u>Education (E)</u>**: "Supercomputing in Plain English" (SiPE) workshop series: 11 talks about advanced computing, taught with stories, analogies and play rather than deep technical jargon. Have reached 362 institutions (academic, government, industry, nonprofit) in 51 US states and territories and 17 other countries (16 OK academic institutions to date).
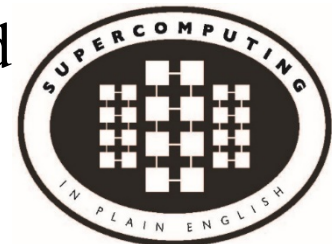
OneOCII Service Methodologies Part 2

- **<u>Faculty/Staff Development (F)</u>**: Workshops held at OU and OSU on advanced computing and computational science topics, sponsored by the National Computational Science Institute, the SC supercomputing conference series, the Linux Clusters Institute, the Virtual School for Computational Science & Engineering. Oklahoma is the only state to have hosted multiple events sponsored by each of these (18 OK academic to date).

- **<u>Outreach (O)</u>**: "Supercomputing in Plain English" (SiPE) overview talk (25 OK academic to date).

- **<u>Proposal Support (P)</u>**: Letters of commitment for access to OneOCII resources; collaborations with OneOCII lead institutions (4 OK academic, 1 nongovernmental).

OneOCII Service Methodologies Part 3

- **<u>Technology (T)</u>**: Got or helped get technology (e.g., network upgrade, mini-supercomputer, hi def video camera for telepresence) for that institution (14 OK academic to date).
- **<u>Workforce Development (W)</u>** – (39 OK academic)
  - Oklahoma Information Technology Mentorship Program (OITMP)
  - "A Day in the Life of an IT Professional" presentations to courses across the full spectrum of higher education.
  - Job shadowing opportunities and direct mentoring of individual students.
  - Institution Types: high schools, career techs, community colleges, regional universities, PhD-granting universities.
- Special effort to reach underrepresented populations: underrepresented minorities, non-PhD-granting, rural

OneOCII Institution Profile

To date, OneOCII has served 103 Oklahoma institutions, agencies and organizations:

- 55 OK academic
- 48 OK non-academic

OneOklahoma Cyberinfrastructure Initiative
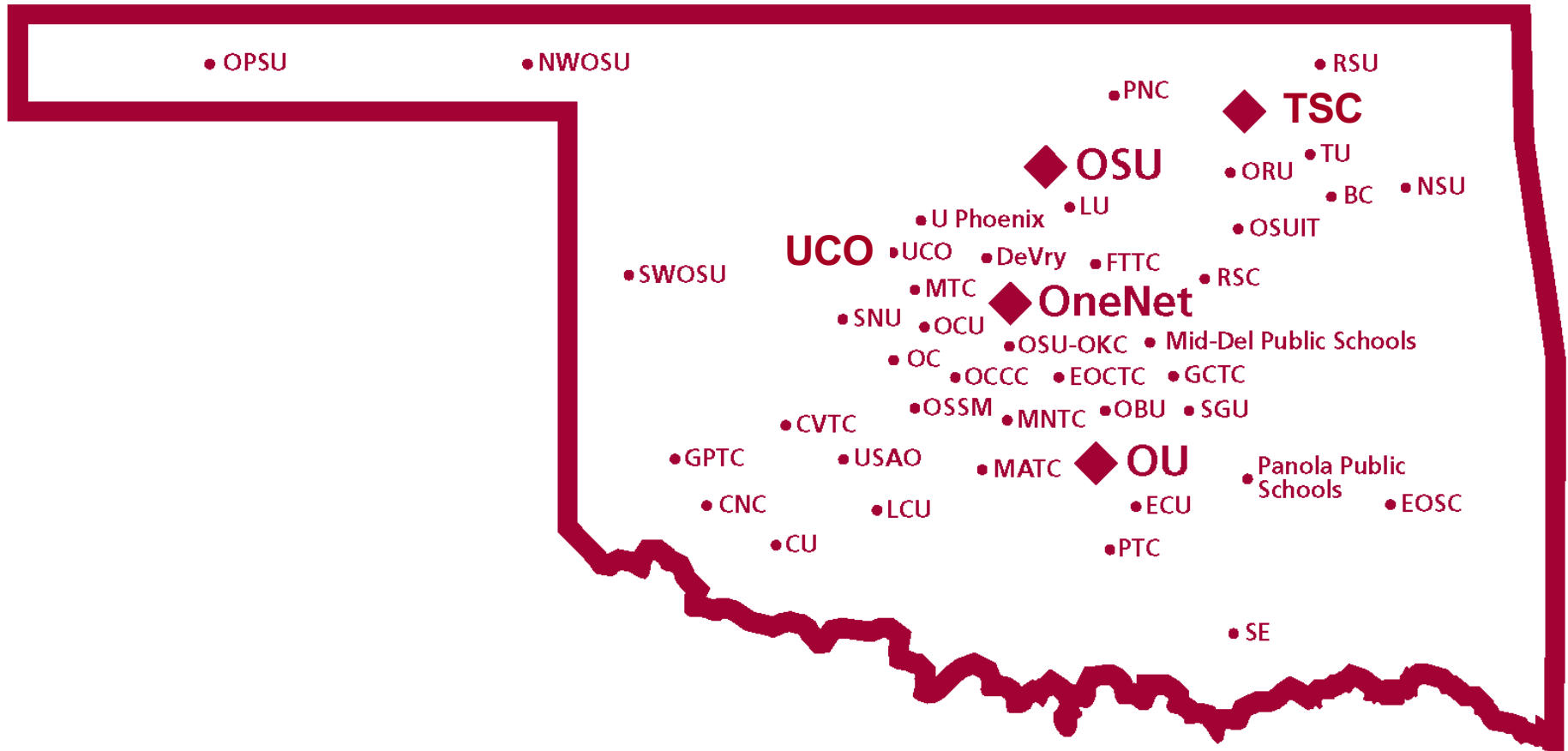
Supercomputing in Plain English

- **FREE and OPEN TO ALL**
- Provided every other spring (coming Spring 2017)
- Available LIVE via videoconferencing
- Topics   http://www.oscer.ou.edu/education/

  1. Overview: What the Heck is Supercomputing?
  2. The Tyranny of the Storage Hierarchy
  3. Instruction Level Parallelism
  4. Stupid Compiler Tricks
  5. Shared Memory Multithreading
  6. Distributed Multiprocessing
  7. Application Types and Parallel Paradigms
  8. Multicore Madness
  9. High Throughput Computing
  10. GPGPU: Number Crunching in Your Graphics Card
  11. Grab Bag: Scientific Libraries, I/O Libraries, Visualization

SiPE Participants

- 362 institutions, firms, agencies and organizations in 48 US states, 3 US territories and 10 other countries
  - Academic: 251 institutions in 51 US states & territories and 6 other countries
    - 69 institutions in 23 EPSCoR jurisdictions
      - 16 institutions in Oklahoma
  - Industry: 49 firms in the US and 4 other countries
  - Government: 44 – US federal and state plus 7 other countries
  - Non-Governmental: 18 (US and 1 other country)
- Missing US states & territories
  - EPSCoR states: RI, VT
  - EPSCoR territories: Guam (no PhD-granting institution)
  - Other territories: American Samoa, Northern Marianas Islands

More States Should Do This

- More states should do intra-state collaboration for Cyberinfrastructure.

- Some states already do similar things, for example:
  - Florida
  - Louisiana
  - New York
  - Ohio

- We want to help other states develop their own.

- We're already scheduled to work with three other states on this in the coming 6 months.

ACI-REF

What is an ACI-REF?

- Advanced Cyberinfrastructure - Research & Education Facilitator

- Expertise in CI and Computational Science & Engineering

- Work with research teams to get them productive and move forward the computational aspects of their research.

Clemson-Led ACI-REF Project

- Original Proposal
  - "The Condo of Condos"
  - 13 institutions, 4 years, $35.7M
  - Included both people (ACI-REFs) and hardware (100GE)
- Funded Grant
  - "Advanced Cyberinfrastructure - Research & Education Facilitators"
  - 6 institutions, 2 years, $5.3M
  - No hardware

Phase 1: Clemson-led ACI-REF

- Phase 1 institutions
  - Clemson U (SC)
  - Harvard U (MA)
  - U Hawaii
  - U Southern California
  - U Utah
  - U Wisconsin Madison
    - All except Harvard are 2012-13 CC-NIE awardees.
    - Only Clemson U and U Hawaii are EPSCoR.
- Of 2012-13 CC-NIE awards:
  - 23% went to institutions in EPSCoR jurisdictions;
  - 46% of EPSCoR jurisdictions got an award (including OK). (Not counting TN and UT, which have graduated EPSCoR.)

Phase 2: Find Other Funding

- Phase 2 institutions
  - Arizona State U
  - Emory U (GA)
  - Ohio Supercomputer Center
  - Stanford U (CA)
  - SSERCA (FL)
  - U Oklahoma
  - U Washington
- OU is the only institution that is all of:
  - EPSCoR
  - 2012-13 CC-NIE awardee
  - ACI-REF Phase 2

OU's ACI-REF Project

OU's NSF CC*IIE Grant

- "A Model for Advanced Cyberinfrastructure Research and Education Facilitators"
- Grant no. ACI-1440783
- NSF CC*IIE program
- 9/15/2014 - 9/14/2016
- $400K
- Totally separate from the Clemson-led ACI-REF grant

OU's ACI-REF Objectives

- Data-Intensive Research Facilitation: Via Software Defined Networking (SDN) across OFFN, facilitate end-to-end management, by researchers, of high bandwidth/high performance data flows through a distributed hierarchy of open standards tools, providing researchers with a new layer of transparency into network transport at OU, among OneOCII institutions, and with ACI-REF members.

- Oklahoma ACI-REF project: Lead and facilitate adoption of the ACI-REF approach across Oklahoma, leveraging extant and emerging capabilities within OneOCII.

OU's ACI-REF Objectives

- <u>National training regime</u>: Provide a "virtual residency" program for Campus CI Engineers and other ACI-REFs, open to not only CC*IIE awardees and ACI-REF members but any institution that needs.

- <u>Research Experiences for Undergraduates (REU) Sites/Supplements</u>: Foster undergraduate research at OU via a culture of integrating REU sites and supplements into Science, Technology, Engineering & Mathematics (STEM) research, including by all research themes on this proposed CC*IIE project.

ACI-REF Role

- Research facilitation: especially in the context of SDN
- Researcher training
- Science DMZ administration
- Statewide participation: facilitating research across the OneOklahoma Cyberinfrastructure Initiative (OneOCII)
- HPC operations

ACI-REF
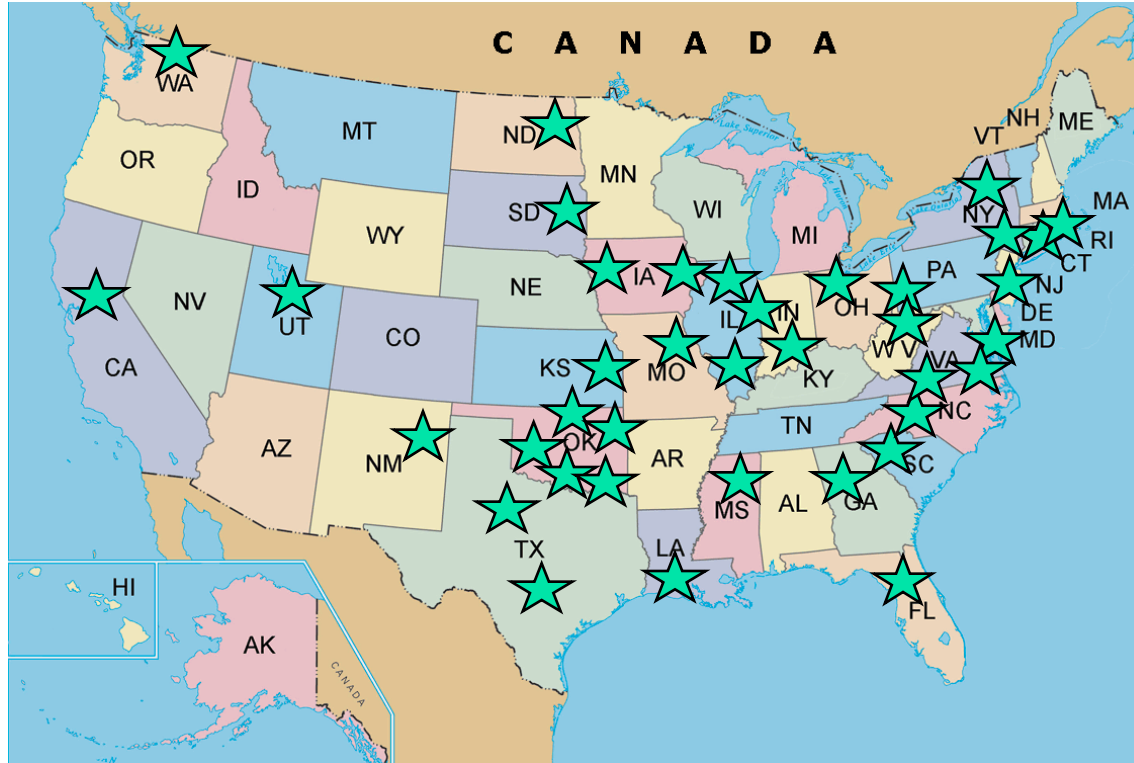# Virtual Residency

ACI-REF Virtual Residency: Why?

- ACI-REFs have strong experience within their discipline (often non-CS).

- Haven't been faculty.

- Sometimes little or no research experience (especially for SDN-focused ACI-REFs).

- Even if strong research background, little or no experience with research outside their own discipline.

- No local, regional or national programs to teach people how to be an ACI-REF.

National Level of Interest

- **Proposal**
  - Polled CC-NIE awardees, MRI CI awardees, Minority Serving Institutions.
  - Interest expressed from 33 institutions in 23 states & territories expressed interest, including 3 MSIs and 19 institutions in 13 EPSCoR jurisdictions, and 7 non-PhD-granting institutions.
- **Applications**: over 60 from 49 institutions in over 30 states and territories.
- **Participants**: 50 total from 39 institutions in 26 states and territories (28 onsite and 22 offsite via videoconferencing), including 21 institutions in 12 EPSCoR jurisdictions, 5 Minority Serving Institutions, and 5 non-PhD-granting institutions.

Virtual Residency Institutions

ACI-REF Workshop Agenda

- SUNDAY (evening pizza party)
    - Welcome and virtual residency overview
    - Introduction to Research Cyberinfrastructure consulting
    - How to Give a CI Tour
- MONDAY
    - Early AM: Effective Communication: How to Talk to Researchers about Their Research
    - Computational and Data-enabled Science & Engineering (CDS&E) Track
        - Mid AM: Deploying Community Codes
        - Early PM: Real user presents their CDS&E research
    - SCIENCE DMZ Track
        - Mid AM: OpenFlow - Lecture
        - Early PM: OpenFlow - Lab
    - Mid PM: CI User Support
- TUESDAY
    - Very Early AM: Project Guidelines
    - Early AM: Faculty: Tenure, Promotion, Reward System
    - CDS&E Track
        - Mid AM: Benchmarking & Tuning
        - Early PM: Real users present CDS&E research
        - Mid PM: Real users: CI consulting practicum ("speed dating")
    - SCIENCE DMZ Track
        - Mid AM: Exploring Open Daylight - Lecture
        - Early PM: Exploring Open Daylight - Lab
        - Mid PM: Real users' high bandwidth research
- WEDNESDAY
    - Early AM: Using Videoconferencing and Collaboration Technologies for Consulting
    - Mid AM: Writing Grant Proposals
    - PM: BREAK (free time)

ONE OCII

OSCER

OU

INFORMATION TECHNOLOGY
The UNIVERSITY of OKLAHOMA

ACI-REF Workshop Agenda

- THURSDAY
    - Early AM: The Shifting Landscape of CI Funding Opportunities
    - CDS&E Track
        - Mid AM: Finding and Provisioning Remote Resources (XSEDE, OSG)
        - Early PM: Real users present CDS&E research ("speed dating")
        - Mid PM: Catch-up on unfinished talks
    - SCIENCE DMZ Track
        - Mid AM: The Software in SDN - Lecture
        - Early PM: The Software in SDN - Lab
        - Mid PM: Real users' high bandwidth research

- FRIDAY
    - Early AM: So You Want to Write a CI Proposal
    - Mid AM: Panel: Stories from the Trenches
    - Early PM: Project work time
    - Mid PM: Project work time
    - Late PM: Project presentations from early departers
- SATURDAY
    - AM: Project presentations

Workforce Meta-Development

Virtual Residency Conference Calls

- Biweekly call via videoconferencing
- All ACI-REF Virtual Residency participants are welcome, as well as some mentors.
- This is where a lot of magic is happening.
- We're trying to figure out how to fund a 2nd cohort.
  - The Virtual Residents have lobbied that we charge them, so that we can put some of their funds into a 2nd cohort.
  - The 1st cohort's 2nd workshop will be the 2nd cohort's 1st workshop.
  - The 1st cohort want to teach and mentor the 2nd cohort.

Find Us at SC15!

- Recruitment Meetings Tue, Wed, Thu
    - Tue Nov 17 2:00-2:30pm Great Plains Network booth # 592
    - Wed Nov 18 1:30-2:00pm Great Plains Network booth # 592
    - Thu Nov 19 10:30-11:00am Indiana U booth # 532
- Virtual Residency All-Hands Meeting
    - Wed Nov 18 2:00-3:00pm booth # 1209

Acknowledgements

Thanks for your attention!
QUESTIONS?