So You Want to Write a Cyberinfrastructure Proposal

Henry Neeman, University of Oklahoma
Director, OU Supercomputing Center for Education & Research (OSCER)
Assistant Vice President, Information Technology - Research Strategy Advisor
Associate Professor, Gallogly College of Engineering
Adjunct Associate Professor, School of Computer Science
ACI-REF Virtual Residency Workshop 2016
Friday August 12 2016

Outline

- Equipment Location and Type (MRI/CRI)
- Vision
- Research Activities to be Enabled
  - Includes Results from Prior NSF Support
- Description of the equipment and Needs
- Impact on Research and Training Infrastructure
- Management Plan
- Cost Share & Institutional Commitment

Read the Solicitation

- The NSF provides an extensive, reasonably detailed solicitation.
  - Read it.
  - Obey it.
- Failure to obey the solicitation can be grounds for returning your proposal without review.

Vision

- Describe the overall project.
- Give your project a name. Use it often.
  - It's great if the name obviously relates to your institution.
- Justify: In your first paragraph, quote a report from a national body (e.g., NSF, national academies, etc) saying that this sort of project is important.
- What will be the outcome if the proposal **isn't** funded?
  - Don't whine about how little your institution has or how old your current infrastructure is. That's not the NSF's problem.
- Give a quick list of research projects.
- Give a quick description of the equipment.
- Provide a table of all projects, including how much of the resource they need, their funding and headcounts.
- If the proposal has no unifying science theme, provide a list of themeless funded grants from past instances of the same program.

Research Projects

- Each research project to be served by the equipment should have anywhere from a page to a few sentences of description, along with the answers to the questions on the next few slides.
  - Prove that it'll be full a lot, ideally oversubscribed.
  - Prove that what fills it is important, **transformative** research.
- Hero projects should get the most proposal real estate and should come first; gradually reduce the length of later, less shiny projects.
- For a Cyberinfrastructure proposal, the biggest consumers of the resource should be early, but you don't have to go in order of consumption -- go from most compelling to least.
- For Small Institution proposals, there's room to use teaching.

Equipment Proposal Questions #1

For each research project:

- How much funding does your research currently have? How much is pending? Planned? From what sources?
- How many faculty, staff, postdocs, grad students and undergrads on your team will be served by this equipment?
- What makes your research transformational?
- What are the broader impacts?
  - Not just in the STEM research sense
    - Underrepresented populations
    - Integration of education and research
    - Dissemination
    - Societal impact
    - etc

Equipment Proposal Questions #2

- How much of the proposed resource (CPU hours, storage, bandwidth, whatever) do you expect to need over the next N years?
- How did you calculate this amount?
- Please give me a one page summary of your research that incorporates these issues.
  - This is typically straightforward, because faculty often have either a 1 page summary from a grant proposal or a more broad research statement.

MRI/CRI for HPC Cluster Questions #1

- How many CPU core hours or node hours will you need over the next N years?
- How did you determine that?
- Have you benchmarked your code?
  - On what platform?
  - What is the expected performance improvement on the proposed equipment, compared to the platform you benchmarked on? How did you come up with that estimate?
  - If it's your own code (or if you're part of the development team for a community code), do you plan to optimize the software? If so, what performance improvement do you anticipate?
    - Be conservative with your estimates. Wild claims of amazing speedup won't impress the reviewers, they'll just think you're naïve or BSing.

http://www.nsf.gov/pubs/2011/nsf11011/nsf11011.jsp

MRI/CRI for HPC Cluster Questions #2

- If the proposal is for a new type of platform (for example, accelerators such as GPUs or Intel Xeon Phi/MIC):
  - Who will be responsible for porting the code to the new platform?
    - If this is either a community code or a commercial code, the porting may already have been done by the developers. Say that.
  - Have they committed to do so? How reliable is that commitment? How is that being funded?
  - What speedup is expected on the new platform? How was that number determined?

MRI/CRI for HPC or Storage Questions

- How much storage will be needed for this project?
  - If this is for live storage: What is the maximum amount of storage at a time that will be needed for this project?
  - If this is for archival storage: What is the total amount of storage needed over the lifetime of the equipment?
- How was that calculated?

CC* Questions #1

- What is the expected typical size of each dataset being transferred?

(It would be helpful to know expected growth rate: Are you expecting it to stay roughly the same over the next several years, or to double every two years, or what?)

CC* Questions #2

- Where are such datasets originating, and where are they being transferred to?

- Why do such datasets need to be transferred between these endpoints?

(That is, what requirement do these data transfers address for your team's research?)

CC* Questions #3

- What is the time window for transferring each such dataset?
- Why does each such dataset need to be transferred during that specific time window?

That is, what's the negative impact of the transfer taking (a) marginally longer and (b) much longer?

- How often do you expect to have such a data transfer need?

Prior Results

- Every NSF proposal has to have a section on “Results from Prior NSF Support.”
- If your team has lots of that, you can’t fit it all. The solicitation and the NSF’s Grant Proposal Guide provide useful guidelines on that.
  - The PI and each Co-PI should each provide the one most relevant grant.
- For an MRI, preference should be given to equipment grants, especially other MRIs.
- If you don’t have anything relevant, say that.
- If you do, is there a way that you can fit this proposal into a more coherent story about how your institution does CI?

Equipment Description

- Give a rationale for the equipment. This reinforces your vision statement.
- Describe the purchase process.
- Describe current similar equipment at your institution or otherwise available to your researchers.
  - Just because there's another one nearby doesn't mean your proposal isn't fundable.
  - Just because there are national centers with bigger resources doesn't mean your proposal isn't fundable.
- Describe the role of the equipment.

Equipment Description

- Describe the equipment in detail. You can express lots of technical information in a modest number of paragraphs.
- Describe how the equipment will be accessed.
- Give the expected performance.
- Describe delivery, installation and operations training.

Impact

- Describe how the equipment will attract researchers. Will it be used to attract new faculty and/or new students?
- Refer back to the research section to describe the impact on your team's research.

Broader Impacts

- Describe the broader impacts.
  - Impact on STEM research
  - Integration of research and education
  - Underrepresented populations
    - minorities, women, disabled
    - non-PhD-granting institutions
    - rural vs urban
  - Societal impact: e.g., health, economic development etc
  - National defense
- Refer back to the individual research projects' broader impacts.
- You don't need to have all of these items, but have some.
- Your broader impacts are judged more on what you've already done that on what you claim you're going to do.

Management Plan

- Describe the facility where the equipment will reside, in detail.
- Operations labor: who, what credentials, who will manage them.
  - You can make them or their boss Senior Personnel.
- Apportioning: Who will get how much?
- Decision making: Describe the procedure.
- Advisory committee(s)
  - External: one CI, one researcher, one broader impacts.
  - You can also have an Internal Advisory Committee.
- Timeline and milestones
- Sustainability plan: What happens when the grant ends?

Management Plan: Facility

- Where will the equipment reside? (e.g., data center)
- What capabilities does that location have?
  - Space
  - Power
    - UPS? Generator?
  - Cooling
  - Fire suppression?
  - Network
  - Security
  - Accessibility: Describe the path from loading dock to data center floor.

Cost Share

- For some programs (e.g., Major Research Instrumentation), cost share is **mandatory** for PhD-granting institutions and non-degree-granting institutions.
  - MRI: 30% of total project budget (NSF pays 70%).
- For some programs (e.g., Major Research Instrumentation), cost share is **forbidden** for non-PhD-granting degree-granting institutions.
- Cost share can only be done at **exactly** the level required in the solicitation.
- There is **NO SUCH THING** as voluntary cost share:
  If they don't ask for it, you can't include it.
  - Your proposal may be returned without review if you do.

Institutional Commitment

- Not the same as cost share.
- Not required, not prohibited.
- Strange rules:
  - **CANNOT** mention any dollar figures (or anything that can be straightforwardly translated into dollar figures).
  - **MUST** appear in the Facilities section (preference for at the end).
  - **SHOULD** be confirmed in a letter of commitment from someone who has the authority to commit.
  - **MAY** appear in the project description (which is a good idea).

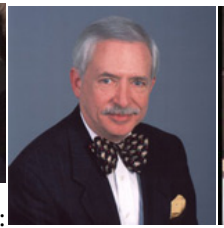
CC: Cyberinfrastructure Plan

- Description of campus network: detail is good!
- Description of statewide network, if appropriate
- Relevant networking grants, if any
- IPv6
- perfSonar
- InCommon
- BCP 38 (or uRPF)
- Connection to Internet2, AL2S, ESnet, etc
- Other relevant services (e.g., Globus Online, Cilogon)

Be concise and specific ….

OK Supercomputing Symposium 2016

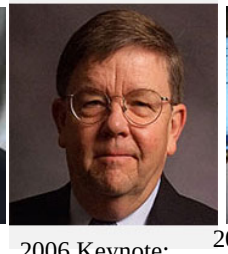
2002 Keynote:
Ron Cooper
Center of Excellence in IT & Telecom
Exec Director



2003 Keynote:
Peter Freeman
NSF
Comp & Info Sci & Engr
Assistant Director

2004 Keynote:
Sangtae Kim
NSF Shared Cyberinfrastructure
Division Director

2005 Keynote:
Walt Brooks
NASA Advanced Supercomputing
Division Director



2006 Keynote:
Dan Atkins
Head of NSF's Office of Cyberinfrastructure

2007 Keynote:
Jay Boisseau
Director
Texas Advanced Computing Center
U. Texas Austin

2008 Keynote:
José Munoz
Deputy Office Dir
Sr Sci Advisor
NSF Office of Cyberinfrastructure

2009 Keynote:
Douglass Post
Chief Scientist
US Dept of Defense
HPC Modernization Program

2010 Keynote
Horst Simon
Deputy Director
Lawrence Berkeley Nat'l Laboratory

2011 Keynote
Barry Schneider
Program Manager
National Science Foundation

2012 Keynote
Thom Dunning
Director
National Center for Supercomputing Applications

2013 Keynote:
John Shalf
Dept Head CS
Lawrence Berkeley Nat'l Lab
CTO, NERSC

2014 Keynote:
Irene Qualters
Division Dir
Advanced Cyberinfrastructure Division, NSF

2015 Keynote:
Jim Kurose
Asst Director
Comp & Info Sci & Engr Directorate, NSF

2016 Keynote:
Dan Stanzione
Exec Director
Texas Advanced Computing Center
U. Texas Austin

FREE!
Wed Sep 21 2016
@ OU

Reception/Poster Session
Tue Sep 20 2015 @ OU
Symposium
Wed Sep 21 2015 @ OU

Thanks for your attention!

Questions?

hneeman@ou.edu