Sustainability for HPC Centers

A Macro View

Jim Bottum, CIO
Clemson University
May 5, 2010
Overview

- NSF history in HPC (*what can we learn from the past*)
  - Program evolution
  - Center evolution
- Thoughts on sustainability
- Review
- Centers in 2020
  - We’re all trying to figure this one out
NSF Program Evolution

• 1970s Federal funding for campus computing recedes
• “Supercomputing Famine” – series of reports in the 1980s (Press, Lax, Bardon-Curtis)
• U.S. response

Office of Advanced Scientific Computing

NSFNET and 5 Centers + NCAR 1985
NSF Program Evolution (cont’)

• Hayes report 1995
  ▫ Recognition of centers “mission creep” at the centers and pervasiveness of CI

• PACI 1997
  ▫ NSF response – declare success and phase out centers,...convert to PACI
  ▫ Mid course correction

• Terascale – 2001
  ▫ Transition from PACI (funding people) to funding capabilities

• OCI - 2005
  ▫ Strategic – director level reporting
  ▫ Balanced set of programs
  ▫ Integration with directorates

• OCI at 5 - 2010
  ▫ Hardware expenditures from 60% of budget 6/06 to over 80% in 12/08
  ▫ Task forces – community assessment; programmatic recommendations
    • HPC, Campus Bridging, Data & Viz, Software, LWD, Grand Challenges & VOs
    • Mid-term reports at ACCI, May 25-26, 2010
### Center Evolution - *general*

<table>
<thead>
<tr>
<th>Year</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>- Small number of university centers&lt;br&gt;- FCCSET (agency knowledge transfer)</td>
</tr>
<tr>
<td>1983</td>
<td>- Phase 1 centers – NSF bought access&lt;br&gt;  - Colorado State, Minnesota, Purdue</td>
</tr>
<tr>
<td>1985</td>
<td>- Phase II – NSF established five centers&lt;br&gt;  - “fully funded”</td>
</tr>
<tr>
<td>1990s</td>
<td>- Emergence of state centers&lt;br&gt;  - 1989 – “several” institutions formed CASC&lt;br&gt;  - WWW, computing pervasive, mission creep</td>
</tr>
<tr>
<td>2000</td>
<td>- Campus centers begingrowth curve as faculty return to comp center&lt;br&gt;  - CASC = 60 members</td>
</tr>
</tbody>
</table>
National Center Evolution - one case

1985
- Two user communities
- No network; 4 processors
- Funding: 90%+ from single source
- Staffing: 75+

1990
- Dozen+ user communities
- Early MPP (R&D); viz; industry; network; software tools
  - Diversification growing
- Funding: 70/30
- Staffing: 175+

1995
- User communities growing
- Early clusters; WWW; Internet2; centers “mission creep”
  - NSF cancels centers; PACI created
- Funding: 50/50; PACI hubs as funding agencies
- Staffing: 400+ (center) PACI partners

2000
- PACI phased out; Terascale - TeraGrid beginning
- Capability computing at centers
- Campuses picking up users in the growing “tail” of users
- Funding: center funding fading; competition for capability
Mission Creep.......one example
HPC Centers 2010 - appear to be three groups

- **National** focus on capability within an evolving framework (11 CASC members)
  - TeraGrid > XD
  - Small # players focused on very high end; competition for federal funds

- **State** – appear to be only 4 in the CASC listing
  - Seems a good model for solvent states
    - Great time to build strategic advantage (Massachusetts)

- **University** focused and based, institutional + research grant support (45 CASC members)
  - Serving some high end users and leveraging resources for other university users
NSF Role in National Networks

...assumption seems to be that the infrastructure job is done or that is someone else’s to do.......
Internet2 Combined Infrastructure Topology
Portfolio of network infrastructure and services across the Internet2 footprint
Facilities-Based Regional Optical Networks

http://paintsquirrel.ucs.indiana.edu/fiber_map.pdf/history.html
That all looks good but still way to go......

Research and Education Capacity

Million bits per second

Facilities-based

Less than 3 years ago

Commodity-based

Clemson

Peers

Top 20

NSF RFP 06-599 High Performance Computing Operations
Clemson could not submit a proposal based on lack of appropriate connectivity
Today........

• Faculty at a comprehensive university in South Carolina have access to less bandwidth than a K-12 student in North Carolina or Florida!
• Faculty drive 4 hours to Clemson because their data takes > 2 weeks on the state network
• Students at a comprehensive that is jointly taught with Clemson are visually limited because they can only turn on 1 of their 3 video cameras at a time due to bandwidth limitations
• Etc.
Thoughts on Sustainability
Things to Consider

- Strategic alignment of mission
- Communicating value and strategy
- Business model
- Evaluation, mid-course correction and “reinvention”
- People
Strategic Alignment of Mission
Institutional Alignment

“Cyberinfrastructure is the primary backbone that ties together innovation in research, instruction, and service to elevate Clemson to the Top 20”
Doris Helms, Clemson Univ. Provost
Communicating Value and Strategy
Investment and Governance Framework

Maslow’s Hierarchy

- Self-Actualization
- Safety
- Physiological

Productivity & Discovery Enabled
Applications, Tools, Consulting
Software Stack/Middleware
Servers, CPU, Storage, ...
Networking/Connectivity
Power & Cooling

Planning, timelines, budgeting, decision-making


Courtesy: John Cobb, ORNL
National High Performance Computing Landscape

Clemson’s Goal: Stay at 10% of the leading edge national centers

- UIUC (NCSA) 2007
  - Sustained PetaFLOPS
  - ~ 2011

- UT Austin (504 TeraFLOPS)
- UT Knoxville (~ 1 PetaFLOPS)
- TBA 2008
- TBC 2009

- Campus centric infrastructure
  - On-ramps

- Cluster Exploratory (CluE)
  - NSF+IBM+Google

NSF Focus FY 06-10

- 5-100 TeraFLOPS
  - Large number of systems

- 100+ TeraFLOPS
  - At least four systems

- Sustained PetaFLOPS at least one system
Clemson CI Days

Over 230 Faculty
Base of 1000

Colleges Attended:
AAH, BBS, E&S, HEHD, AFLS & Libraries

Dean of Arts & Humanities chaired a panel

National Panel & CNI
Business Model
Diversification.....

.......*how many legs are on your stool?*
Why Diversify?

CCIT Funding Sources

- Core E&G
- Student Fees
- Services - Auxiliary - Self Gen
- Grants, Gifts, Contracts

FY11
Clemson started out...
...and is morphing into...Cyberinstitute

- **Testing & evaluating technologies & services**
  - CoEE Endowed Chair in Cyberinfrastructure ($4M)
    - Home in Electrical and Computer Engineering
  - Department of Energy ($2.5M)
    - South Carolina “Cloud Testbed”
  - National Science Foundation Research Information Infrastructure Track 2 ($6M)
    - Clemson leads partnership between SC and Tennessee
  - Industrial partnerships being developed

- **South Carolina Cloud:**

  - 4 universities, 1 high school
  - STEM Teaching – Parallel Computing – Grid Classroom
  - HPC remote access & remote system administration
  - application hosting
  - DR and Business Continuation

- **Industrial partnerships being developed**
Technology Transfer

Royalties

- General Education
- Family & Community Living
- Biotechnology & Biomedical Sciences
- Advanced Materials
- Information & Communication Technologies

We’re in the technology business….and institutional enlightenment can help!
Evaluation, Mid-course Correction and Reinvention
REPORT OF THE ADVISORY COMMITTEE FOR GPRA PERFORMANCE ASSESSMENT

FY 2009

DAVID B. SPENCER, Sc.D.  
CHAIR

SHARON DAWES, Ph.D.  
VICE CHAIR


What is your evaluation framework?

NSF showing results to the Hill
How do we get customer feedback....

....and how often?
.....broader impact

- **Architecture students using Condor Render Farm**
  - Up to **three hundred** graduate and undergraduate students
  - All students regularly use computer aided drafting tools
    - Autodesk VIZ, Maya, and Rhino
  - Rendering large scenes with multiple light sources
  - Capability allows larger projects to get done in less time
    - have to “dumb down” the project if done on laptop or lab computer

..... which side of the bread is the butter on….HPC v HTC – who cares!
Being Nimble

- Good v bad money
  - Marketing reaction
  - “Fat” dollars are the best

- Becoming very good at something that no one wants anymore
  - May have to break up or kill something that was technically very successful but demand is no longer there

- Reinvention and even adaptation can be painful
  - In our business it happens all the time

- Headroom is critical to nimbleness
  - And with the “new normal” it is harder than ever to achieve
  - Does this state push us toward increased collaboration?
People

• Are people working for a cause, a paycheck or because they want to live in a particular place
  ▫ 10 years – 2 stories

• Morale
  ▫ Seems to be highest when vision is clear, staff’s work can be linked to the vision and is clearly defined and resources are there to get the job done (Burnham)

• Challenge....development....support
  ▫ Are we providing; setting example; clearing the path?

• Users – customers.......partners
  ▫ Are we in touch......do we stay in touch.....
  ▫ Forum

• Pipeline - students
  • Are we involving students enough?
  • Are they truly being trained on architecture and systems ie, under the hood?
Summary
Summary

• **1970s - campus based computing was prominent**
  - At least 2 things changed this (VAX and Nixon)

• **1980s – NSF stimulated university based HPC**
  - Centers were successful and emulated
  - Many universities stopped investing because centers were there

• **1990s – technology permeated and proliferated**
  - Networks and WWW changed the world as we knew it and the world became flat
  - Technology, security and data deluge begins driving faculty back to campus data centers

• **2000s**
  - Campus CIOs now a critical part of the high end computing environment
  - “Great recession” forces increased need for an entrepreneurial approach to sustainability of growing base of HPC centers
  - Very likely that collaborative creative partnerships across states and regions will be necessary for high end sustainability
Centers in 2020
Common Center Characteristics

• Recognized differentiator
  ▫ Will we be smart enough to break down our silos and leverage others’ strengths?

• Enlightened, innovative governance model
  ▫ Cross organizational boundaries

• Diversification of revenue
  ▫ Tap root but also other features: revenue sharing?

• Clear vision, focus and commitment to excellence
  ▫ Daniel Burnham
June 11, 2009 - For immediate release:

PATRICK ADMINISTRATION ANNOUNCES COLLABORATIVE PLAN TO BUILD HIGH-PERFORMANCE COMPUTING CENTER & RESEARCH PROGRAM IN HOLYOKE.

Unprecedented collaboration with MIT, UMASS, BU, EMC & Cisco Systems will position Massachusetts as leader in next generation of computing technologies

HOLYOKE – Thursday, June 11, 2009 – Governor Deval Patrick, MIT President Dr. Susan Hockfield, UMASS President Dr. Jack M. Wilson, Boston University President Dr. Robert Brown, Joseph M. Tucci, Chairman, President & CEO of EMC, and John T. Chambers, Chairman & CEO of Cisco Systems today announced a collaboration that will lead to the development of a world-class, high-performance computing center in Holyoke, and a statewide collaborative research program. The center will be powered by a combination of green and cost-competitive energy, making it a cost-effective and environmentally sound facility.

Today’s announcement positions Massachusetts to be a leader in the application and development of the next generation of computing technologies, and will spur the creation of an innovation economy district in Holyoke.

“Innovation is our calling card in Massachusetts, and this partnership with MIT, the University of Massachusetts, Boston University, EMC and Cisco Systems will usher in the next generation of high-performance computing and set us apart from our competitors,” said Governor Deval Patrick. “The potential for breakthrough technologies and research is enormous, and both the center and this collaboration will undoubtedly serve to lift up the City of Holyoke and regional economies throughout Western Massachusetts.”

High-performance computing has become an increasingly critical tool for both academic and corporate research across a variety of fields, including biotechnology and alternative energy. The commitment of the center’s partners to collaborate with state, academic, and business institutions across industries will create jobs in Western Massachusetts and power scientific and technological research to ignite regional economic growth.

“Many of today’s most important technical challenges will yield only to the power of high-
Challenges

• “Centeritis” – top of funding pyramid may be smaller than 15 years ago
• Real leveraged collaboration – remember the “metacenter” and all the disincentives
• Cloud dependency vs. innovation using clouds as a basis
• Networks that track capacity change in data/computing
• Strong headwinds against networking at state level from the telcos
• How does the community stay in front not behind?
The End

....is just another beginning