Social Networks, Cyberinfrastructure (CI) and Meta CI

Daniel E. Atkins
School of Information & Dept. of Elec. Engr. & Comp. Sci.
University of Michigan
Ann Arbor
atkins@umich.edu
November 2005
A Little Reminiscing

circa 1965
Illiac I (1949-1962)

Illiac I was the first computer built and owned entirely by a university. It went on line on September 22, 1952. With the computing power of a modern-day handheld calculator, Illiac had 2,800 vacuum tubes and weighed 10,000 pounds. It was more than 10 ft long, 2 ft high, and 8 ft high, with a 5k main memory and 64k Drum memory.
Illiac 2 (1956-68)

One of the first transistor machines. Each transistor hand tested and cost $80 each.
Illiac III (1960-68)

- The ILLIAC III was a fine-grained SIMD pattern recognition computer built by the University of Illinois in 1966.
- This ILLIAC's initial task was image processing of bubble chamber experiments used to detect nuclear particles. Later it was used on biological images.
- The machine was destroyed in a fire, caused by a Variac shorting on one of the wooden-top benches, in 1968. (Including my four floating point AUs.)

From Wikipedia
Illiac IV (1965-82)

Vietnam Era, highly parallel GigaFLOPS machine
See Display in CS Dept. Next Door
Overview

• The CI Movement
• Recent Developments at NSF
• CI Duality w.r.t. Social Network Research
• The Meta-CI Challenge
• Call for Help from the SNAC community.
“a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today’s challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive “cyberinfrastructure” on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy.”

http://www.nsf.gov/od/oci/reports/toc.jsp
Dozens of Workshops on CI and Discipline X

Several internal NSF studies

CISE Div. of Shared CI

Office of CI (OCI) and the CI Council

Draft 4 of NSF CI Vision for 21st Century Discovery & request for input.

RFP: High Performance Computing System Acquisition: Towards a Petascale Computing Environment for Science and Engineering

RFP: Cyberinfrastructure for Environmental Observatories: Prototype Systems to Address Cross-Cutting Needs

See NSF website for more details.

ACLS-Mellon Cyberinfrastructure for Humanities

HASTAC Project
NSF states intent to “play a leadership role”

- NSF will play a leadership role in the development and support of a comprehensive cyberinfrastructure essential to 21st century advances in science and engineering research and education.
- NSF is the only agency within the U.S. government that funds research and education across all disciplines of science and engineering. ... Thus, it is strategically placed to leverage, coordinate and transition cyberinfrastructure advances in one field to all fields of research.

From NSF Cyberinfrastructure Vision for the 21st Century Discovery
Cyberinfrastructure-enhanced Knowledge Communities (Networks)

Outcomes: New Ideas, New Tools, Education & Career Development, Outreach*

Attributes: Collaborative, Multidisciplinary, Geographically Distributed, Inter-institutional*

Specific Cyber Environments:
collaboratories, grids, e-science community, virtual teams, community portal, ...

Cyber-infrastructure Services
Equipment, Software, People, Institutions

Computation, Storage, Communication and Interface Technologies


Broader Application to other disciplines and types of activity.
From CI Advisory Panel Report

Community-Specific Knowledge Environments for Research and Education  
(collaboratory, co-laboratory, grid community, e-science community, virtual community)

Customization for discipline- and project-specific applications

<table>
<thead>
<tr>
<th>High performance computation services</th>
<th>Data, information, knowledge management services</th>
<th>Observation, measurement, fabrication services</th>
<th>Interfaces, visualization services</th>
<th>Collaboration services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Networking, Operating Systems, Middleware</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Base Technology: computation, storage, communication

= cyberinfrastructure: hardware, software, services, personnel, organizations

towards functionally complete CKCs
Cyberinfrastructure includes both

- **Technology Infrastructure** (creation and provisioning) - middleware, portals, HPC, hybrid (IP & lambda) networks, ...

- **Social Infrastructure** (competition & cooperation, IP policies, incentive structure, cost, etc.)
Recent NSF CI and the Social Sciences Workshop

http://vis.sdsc.edu/sbe/

- Six sub-themes
  - CI-mediated interaction
  - CI tools
  - Organization of CI and CI-enabled organizations
  - Economics of CI
  - Malevolent uses of CI
  - Impact of CI on jobs and income
From NSF Cyberinfrastructure Vision for 21st Century Discovery

4. Education and Workforce

3. Collaboratories, observatories, virtual organizations

“sophisticated” science application software

1. Distributed, scalable up to petaFLOPS HPC
2. Data, data analysis, visualization

- includes networking, middleware, systems software?
- includes data to and from instruments?

- provide sustainable and evolving CI that is secure, efficient, reliable, accessible, usable, and interoperable
- provide access to world-class tools and services
HPC Strategy

System Diversity

1-10 Peta FLOPS
1-50 TeraFLOPS
100+ TeraFLOPS

At least one system
Multiple systems
Significant number of systems

Leading Edge Level
"supports a more limited number of projects with highest performance demand"

National Level
"supports thousands"

Campus Level

On demand & dynamic resource allocation.
HPC in a relevant-time loop.

D. E. Atkins
atkins@umich.edu

NSF Focus FY 2006-10

100+ TeraFLOPS
1-10 Peta FLOPS
1-50 TeraFLOPS

Cluster
Grid

Single chip processors with various internal architectures
Our panel proposed a 3 component, CI program with potential to revolutionize the conduct of science & engineering research and applied education.

R&D for technical and social system architecture

Creation and provisioning of CI

Transformative (Revolutionary) use within research & allied education communities

Borromean Ring Synergy
Assertion of “revolution” is based on assumption of

- Significant new, long-term funding.
- Coordinated investment in all three components of activity. (B-Ring synergy)
- High leverage through coordinated co-investment by many stakeholders in addition to NSF.
And while we are at it...

• can we create CI environments in support of research, learning, and societal engagement in ways that exploit complementarity between them?

• Pasteur’s Quadrant research
• Ubiquitous learning environments
• Authentic learning
• Professional development

Knowledge Communities* Enabled by Cyberinfrastructure

The CLEAR Agenda
The OPEN CLEAR Agenda

The Openness Movement
Cl Duality

• Cl as an object of research. “Science on Behalf of Science”

• Cl as a means of research.

• Cl as a object of social networks research

  • Task 1: The identification of ways in which theoretical and methodological insights from recent advances in social network analysis can inform the development of cyberinfrastructure to enable networks within distributed communities of scientists, educators, and practitioners. The workshop will acquaint the scientific and engineering community with the potential of social networking tools to enhance communication and knowledge networks within communities supported by cyberinfrastructure.

  • Task 2: The identification of ways in which cyberinfrastructure can help the community of social network researchers address grand challenges in network science. The workshop will assess the viability and contours of developing a Social Network Analysis Cyberinfrastructure (SNAC) for advancing network analysis, modeling, visualization research, practice, and education.
CI as the *means for* social networks research (SNR)

- Use modelling/simulation to find what patterns of emergent behavior occur in models of very large societies?
- Instrument CI environments to be living labs for SNR research.
- Create global collaboratories for SNR including community databases and tools.

* From NSF Vision 21 Report
CI as an object of social networks research

- What strategies might be developed to optimize management of complex infrastructure systems?*

- How does major technological change affect human behavior and structure complex social relationships?*

- Instrument CI environments to find principles of design for and assessment of impact of CI environments. Help document tangible advantages of knowledge nets enabled by CKCs. Make the advantages more explicit.

- Explore conditions of diversity, independence, decentralization (aggregation) for “wisdom of crowds” phenomena in CKCs. (ref. is to James Surowiecki book)

- Participate in both the analysis and the design of CKCs.

* From NSF Vision 21 Report
Are there fruitful links between

• the Social Networks research community and

• the emergent Incentive Centered Design (ICD) Community?

• see http://www.socialcomputingresearch.net/twiki/bin/view/ICD/WebHome

• We believe that careful attention to individual incentives can lead to vast improvements in the design of systems that rely on information, communication and collaboration technologies to mediate interactions. We draw on theories of rational decision making, game-theoretic models of strategic interaction, and economic, psychological and social theories of motivation to understand the likely behaviors of individuals in response to various system configurations. Based on that understanding, we create or identify design options that meet various normative criteria, such as efficiency and fairness. One hallmark of our approach is that we seek to have an impact on designs of real systems—our projects typically are inspired by a practical problem, move into the realm of abstract theorizing, and end by influencing the design of fielded systems.
Summary so far...

- Mounting a potentially revolutionizing, advanced cyberinfrastructure program is very complex and will not emerge solely as a consequence of technological determinism. Real initiative, new resources, new social structures, and leadership are required.

- It requires the nurture and the synergistic alignment of three types of activity:
  - research and development on the technical and social architectures of CI-enabled science;
  - reliable, evolving, and persistent provisioning of the CI services; and
  - transformative use through iterative adoption and evaluation of CI services within science communities.

- All this should be done in ways that extract and exploit commonality, interoperability, economies of scale, and best practices at the CI layer.

- It will also require shared vision and collective action between many stakeholders including research funding agencies, universities, and industry.

- An even bigger challenge and opportunity is to mount CI programs in ways that benefit and connect and leverage research, learning/teaching, and societal engagement at all levels of education, and in a broad range of areas including the arts and humanities.
The META CI Challenge
What (new) social structures and incentive systems do we need to make the CI movement successful (revolutionary)?

What is the infrastructure for the cyberinfrastructure movement?

Here are some tries at talking about the challenges.
A partial list of complex attributes....

- **Institutionalized providers** of ICT-based tools and services together with expert knowledge and training that serves many specific, customized knowledge communities (overlapping, interoperable, open).
- Architecture and processes that **identify and exploit commonality**, and accommodate heterogeneity through middleware and open standards.
- **Shared creation and re-use** of software, information, facilities, and best practices to promote cost-effectiveness and efficiency.
- **Tight coupling with R&D** in CISE, appropriate SBE areas, and pioneering application areas.
- Sustained commitment, evolutionary enhancement, and "career worthiness"
- Appropriately trained **human resources** for the creation, provision, and application of CI. (blend of ICT, sociology, disciplinary expertise)
- Support for **integrated socio-technical evaluation** to understand impact and inform iterative design processes.
The social challenges in the NSF CI Vision 21 Language

- lead
- leverage
- coordinate investment and activities
- engage in and benefit from
- transition advances (field to field)
- partner with key stakeholders (national and international)
- broaden participation; strengthen workforce
- inclusive strategic planning
- integrate resources and activities
- share resources
- establish collaborative governance structure
- develop human-centered CI
Balanced investment in enhancement of technology capacity and “completeness” of CI environments

usefulness of the cyber environment
A system of cooperative activities:

- Cooperative activity within the CISE basic research community on the fundamentals of software, networking, computation, and HCI enabling cyberinfrastructure.
- Cooperative activity within the research university community to understand and invest in CI to serve S-E research as well as the future of its complete mission.
- Cooperative activity to build shared vision, excitement, supportive policy and co-investments by Federal and State governments, universities, and industry to create, operate and sustain cyberinfrastructure.
- Cooperative activity between science-driven CKCs, and the SMET learning R&D community to explore dual-use learning opportunities.
- Cooperative activity between the interested subset of the CISE research community and the requirements and evaluation testbeds provided by the cooperating e-science projects.
- Cooperative activity between existing science-driven CKC* (e-science) projects to cross pollinate and find common needs and solutions. Use the passion, reputation, experience of the pioneer leaders and their projects.
- Cooperative activity between CISE research community, advanced applications and industry to create open standards/open source CI software.
- Cooperative activity between NSF, LC, other Fed Agencies, academic libraries, and others on R&D and provisioning of curated, persistent, interoperable data archives.
- Cooperative activity between NSF, DOE, DOD and vendors on R&D and provisioning of highest end computing & networking.

CKC = cyberinfrastructure-enhanced knowledge communities
Alignment of stakeholders towards achieving strategic goals

STRATEGIC GOALS
- Significant advancement in discovery, learning, and participation in science and engineering.
- Enhanced innovation and competitiveness in a global, knowledge-based economy.
- Collateral benefits for non-science disciplines, and the overall mission of universities.
- Enhanced human well-being.

ACTIVITY TYPES
- R&D for technical and social system architecture
- Sustained provisioning of advanced cyberinfrastructure
- Transformative use within research & allied education communities
- Creation/ modification of supportive policy

STAKEHOLDERS
(U.S. & other nations)
- Researchers: individuals, projects, disciplinary communities
- Research and Educational Institutions:
  Universities & Colleges
  Government and Private Labs
- Funders: NSF, other Federal agencies, private foundations, State governments, universities, industry
- Infrastructure Providers: Industry, non-profit enterprise
So what can

- the social network research community,
- you and your institutions, and
- this workshop

DO

- to not only explore the duality of SNR and CI, but also
- to help define and create the social structures in which the CI movement will prosper and have broad, positive impact on the world.
Questions, Comments, Discussion