Mythbusting
Scientific Knowledge Transfer with nanoHUB.org
Gerhard Klimeck, Purdue University, gekco@purdue.edu

Over 12,000 / 230,000 Users Annually

nanoHUB.org usage 2012-02-03 00:00:00
Thanks to

nanoHUB contributors:  
330+ tool authors  
1,000+ content authors

Research Group  
@Purdue  
@NASA JPL 1998-2003  
@Texas Instruments 1994-1998

nanoHUB and HUBzero Team
1965
Gordon Moore

http://www.intel.com/technology/mooreslaw

Relative Manufacturing Cost per Component

Number of Components per Integrated Circuit
Intel in 2009

Device Size:
Tens of


Device Integration:
>2 Billion

Stanford SUPREM

Berkeley SPICE

RoberChau(Intel), 2004
Berkeley Simulation Program with Integrated Circuit Emphasis.

from: Larry Nagel, BCTM ‘96

• Started as a class project
• Developed as a teaching tool
• Quality control: pass Pederson
• Dissemination:
  ▶ Public domain code
  ▶ Pederson carried tapes along
  ▶ Students took it along to industry and academia
  ▶ Released 1972
Stanford University Process Modeling

- Stanford wanted to mimic Berkeley success
- Combine various existing models
- Dissemination:
  - Public domain code
  - Community workshops
  - Students took it along to industry and academia
Birth of an Industry

Intel Capitalization: $85B
Total Industry: $280B
What’s Next?

Nano Initiatives

Device Size

- nano-scale structures

Years

- Billions of nano structures

Transistors

Research

- Electronics
- Materials
- Photonics
- Mechanics
- Bio/Medicine

What's Next?
Extensive Facilities
Nano Models

Carbon Nanotubes

Quantum Dots

Artificial Atoms
Computational Nano
Computational Nano
Different Worlds
Imagine Breaking Barriers

Easy use
No Install
Any Browser
Industries Breaking Barriers

Easy use
No Install
Any Browser

Imagine Breaking Barriers

Why is this so hard?
Most research codes are written by one user for one user.
Why is this so hard? Most research codes are written by one user for one user?

Structure
{
  Material
  {
    name = GaAs
    tag = substrate
    crystal_structure = simplecubic
    atoms = (GaAs)
    Lattice:a_lattice = 0.565
    regions = (1)
    Bands:TB:s:param_set = nanohub
    Bands:TB:s:nanohub:E_S_GaAs = 12.1307935176
    Bands:TB:s:nanohub:V_S_S_Sigma_GaAs_H = -20
    Bands:TB:s:nanohub:passivation_potential = 125
  }
  Domain
  {
    name = structure1
    type = pseudomorphic
    base_material = substrate
    dimension = (18.0,19.0,9.0)
    periodic = (false, false, false)
    crystal_direction1 = (1,0,0)
    crystal_direction2 = (0,1,0)
    crystal_direction3 = (0,0,1)
    space_orientation_dir1 = (1,0,0)
    space_orientation_dir2 = (0,1,0)
    regions = (1)
    geometry_description = simple_shapes
  }
}
User Hostile
Why is this so hard?

Most research codes are written by one user for one user?
Why is this so hard?

Most research codes are written by one user for one user.

Accessible (no installation)

Developer Friendly

User Friendly

HubZero

Rappture
It has been very hard!

Emerged Myths

Accessible (no installation)

Developer Friendly

User Friendly

HUBzero

Rappture
Emerged Myths

User Friendly
Cannot use research codes for education
Must write own code to do research
Experimentalists cannot use research codes

Developer Friendly
Building User Interfaces too Difficult
Must rewrite code for web deployment
There is no incentive to share codes

Accessible (no installation)
NO End-to-end Science Cloud Possible
Usual Science Gateway Process

- 175 tools / 4 years: => $88M
- $500k/tool
- NO new research!
- Not validated by researcher (disowned)
- Researcher has much better version
- Code rewrite takes 2-3 years

Many Proposals read alike
Usual Science Gateway Process

- 175 tools / 4 years: => $88M
- $500k/tool
- NO new research!
- Scale back expectations
- Not research codes
- Toy applications
- Not deep research
- Maybe for education?

Generating a Bad Reputation
nanoHUB Process

- 175 tools / 4 years without $88M
- Eliminate bottlenecks
  - No Middleman
  - No Rewrite
  - Retain ownership
- Rapid Deployment: 2-3 years → 1-2 weeks
- Rappture toolkit
- S/W Dev. Ecosystem

nanoHUB is different
Software Development Ecosystem

Development Workspace

Repository

Staging & Testing

Researcher

NanoHUB.org – nowhere else!

Tool Projects

Team

Repository

Wiki Notes

User Support

Q&A

Tickets
Continual Engagement

v1

v2

v3

nanoHUB can prove it

Over 300 Developers

NOT PAID by NCN

Cumulative

- Tool Versions
- New Tools
- Active Developers

Introduction of web-based developer training

Continual Engagement

Over 300 Developers

NOT PAID by NCN
Cyberinfrastructure for Developing Tools

Web-based Publishing System

What’s Next?

Your latest code is installed and ready on nanoHUB.org. Please test your tool by clicking the button below to make sure that everything is working properly, as well as verify that the page describing your tool is created and displays correct information:

- Test your application: Launch tool
- Review the page describing your tool

We’re waiting for You

Once you tested your tool and verified that it is working properly, click here to let us know:

- My tool is working properly, I approve it.

Need to make changes? Once you’ve checked in your latest fixes, click here to let us know:

- I’ve fixed my code. Please install the latest updates.
Software Updates and Deployment

January 1, 2009

>200 S/W projects
>300 developers
Managed by ONE person!
Good news!
In just 5 years, nanoHUB.org hosts
190 new tools plus
1000 version updates
Good news!
In just 5 years, nanoHUB.org hosts
190 new tools plus
1000 version updates

Bad news!
In 1 year, 1/3 tools are new or updated
In 2 years, 2/3 tools are new or updated
– must manage this
Server-side tools have inherent lifecycle management

Current versions served by default – problem solved
Cyberinfrastructure for Running Tools

- HUBzero Technology Group
- Purdue University

Linux/Apache/MySQL/PHP

Physical Machine

Virtual Machine

Content Database

Maxwell’s Daemon Middleware

Open Science Grid

DiaGrid

TeraGrid™

nanoVIS

tool session cluster

Rendering Farm

nanowire job
Building Interfaces and Data Management Systems Fast

Rappture

MATLAB®

Fortran

F77

Python

Perl

Ruby

Scientist

Simulation Code

• Rapid Application Infrastructure
• Created by NCN in Nov 2004
• Open Source (rappture.org)
• Create standard desktop apps
• Works with your favorite programming language

Open Source (rappture.org)
Developer Collaboration Network

White dots: outside NCN

Each dot is a tool
Links are people
Developer Collaboration Impact

Users Served vs. Collaborators

Orange dots: Site Leads
White dots: Developers
Next Generation Publications
Research Incentives
Tool Usage \approx reading papers

17 tools
\rightarrow 11,570 users
\rightarrow 123 citations
Recently Dr. Ahmed was promoted to tenured Associate Professor. I would like to emphasize that Dr. Ahmed's use of nanoHUB in education and research, which earned him national and international visibility, did play a significant positive role in his early promotion case.

Glafkos Galanos
Chair, Dept. of Electr. and Comp. Eng, SIUC
nanoHUB on iTunes U

Nov 2009 start
350 content items today
55,000 downloads
~10,000 downloads/month
Wikipedia Contributions

16 animations deployed Jan 2010 on ~30 pages
Brings 2,200 visitors for 4,000 visits monthly
Building User Interfaces too Difficult

Must rewrite code for web deployment

There is no incentive to share codes

Myths Busted

Developer Friendly

260+ Interactive Tools

Over 300 Developers (mostly volunteers)

Rappture

HUBzero
Emerged Myths

Activities on http://nanoHUB.org in 172 countries

- New Registrations
- Simulation Users
- Tutorial / Lecture Users

nanoHUB.org usage 2012-02-03 00:00:00

Cannot use research codes for education

Accessible (no installation)

Must write own code to do research

NO End-to-end Science Cloud Possible

Experimentalists cannot use research codes
User Behavior Analysis

[Diagram showing data visualization with time (days) and users]
Formal Education vs. Research

A  Soph. Materials Engineering
B  Soph. Mechanical Engineering
C  Senior Electrical Engineering
D  Freshman Chemistry
E  Graduate Electrical Eng.
F  Soph. Materials Engineering
G  Experimentalist Researchers
H  Computational Researchers
I  Self-Study Users

July 1, 2009  Time (Days)  June 30, 2010
**Formal Education vs. Research**

**AY 09/10:**
- 116 Courses,
- 97 institutions,
- ~2,100 students
- 95% outside NCN

**Myth Busted:**
Proof of use in EDUCATION!

**Knowledge Transfer out of Research**
- 14,521 students,
- 761 courses,
- 189 institutions

**Tools and Usage Pattern**

<table>
<thead>
<tr>
<th>Tools and Usage Pattern</th>
<th>Valid Users Known</th>
<th>Classes Like This</th>
<th>Total Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td>1392</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td>253</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>84</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td>803</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td>142</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td>94</td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td>5,685</td>
</tr>
</tbody>
</table>

Voluntary / Viral Use
Time to First Adoption

Typical textbook update: 3.8 years

Time Between Tool Publications and First Use in Classroom
Rapid Adoption of Research

Median adoption time: 174 days (5.7 months)

Typical textbook update: 3.8 years

Time Between Tool Publications and First Use in Classroom
Median adoption time: 174 days (5.7 months)

Typical textbook update: 3.8 years

Time Between Tool Publications and First Use in Classroom
User Behavior Analysis

=> Is Research Possible?

- single tool, single use / homework
- single tool, sustained, intense
- multiple tools, sustained, periodic
- multiple tools, small class
- research, self-study

Over 1,660 authors, 77% non NCN
Proof of voluntary use by OTHERS
Myth Busted
Proof of use in RESEARCH!
Myth Busted: Proof of use in RESEARCH!
Over 1,660 authors, 77% non NCN
Proof of voluntary use by OTHERS
56% outside NCN

857 nanoHUB Citations
each dot is a paper line is common author
Myth Busted: Proof of use in Experimental Work! Not just computational theory!

Barrier Broken 9% of papers by Industry Authors
Academy of Engineering Member

Faculty member 3 years after PhD

h-index: Research Quality Indicator

Year 2009
575 nanoHUB citations
>3,200 secondary citations
h-index: 27

Year 2012
857 nanoHUB citations
>7,500 secondary citations
h-index: 41

"deep research =? many CPUs"
Compute Intensive: NEMO/OMEN

18 years development

- Texas Instruments
- NASA JPL
18 years development

- Texas Instruments
- NASA JPL
- Purdue
Compute Intensive: NEMO/OMEN

18 years development

- Texas Instruments
- NASA JPL
- Purdue
- Peta-scale Engineering

ACM Gordon Bell Prize
Honorable Mention

Mathieu Luisier, Timothy B. Boykin, Gerhard Klimeck, Wolfgang Fichtner

Atomistic Nanoelectronic Device Engineering with Sustained Performances up to 1.44 PFlop/s
• Science, Nature Nano

Ohm's Law Survives to the Atomic Scale

B. Weber,¹ S. Mahapatra,¹ H. Ryu,²* S. Lee,² W. C. T. Lee,³ G. Klimeck,² L. C. L. Hollenber

As silicon electronics approaches the atomic scale, shrinking in size to the active device components. Maintaining its scale is challenging because of the presence of confining surfaces and interfaces. We report on the

A single-atom transistor

Martin Fuechsle¹, Jill A. Miwa¹, Sudhasatta Mahapatra¹, Oliver Warschkow¹, Lloyd C. L. Hollenber², Gerhard Klimeck²

AcM Gordon Bell Prize
Honorable Mention

Mathieu Luisier, Timothy B. Beykin, Gerhard Klimeck, Wolfgang Fichtner

Peta-scale Engineering

Gordon Bell

Texas Instruments
NASA JPL
Purdue

Network for Computational Nanotechnology

Compute Intensive: NEMO/OMEN
18 years development

- Texas Instruments
- NASA JPL
- Purdue
- Peta-scale Engineering
- Gordon Bell

Powers 8 Tools:
10,837 Users
166,793 Simulation Runs
Myth Busted! Prove Computational Extensive Work

18 years development
- Texas Instruments
- NASA
- Purdue
- Peta
- Gordon
- SciPy

267,362 Simulation Runs
10,786 Users
3,874 Users in 100 classes

166,458 Simulation Runs
Resource Requirements
Simulations vs CPU Consumption
Resource Requirements
Simulations vs CPU Consumption
Resource Requirements
Simulations vs CPU Consumption

Classroom Behavior
each dot one user

Research Behavior
each dot one user

Used CPU Time

Number of Simulation Runs

Class A
Class B
Class C
Class D
Class E
Class F
All Classes

Group H: Comp. Res
Group I: Self-Study

Used CPU Time

Number of Simulation Runs

1 10 10² 10³ 10⁴ 10⁵
Resource Requirements
Simulations vs CPU Consumption

Research:
• Avg: 200 runs, 8 hours CPU
• Top: 10,000 runs, 10,000 hours

Education:
• Avg: 20 runs, 5 minutes
• Top: 400 runs, 20 hours CPU
Usage Patterns

=> Tool Qualification

Each dot is one tool
Size of dot indicates number of users

Tools Ranked by Frequent Use in Teaching
Usage Patterns

=> Tool Qualification

Each dot is one tool
Size of dot indicates number of users

Tools Ranked by Frequent Use in Teaching
Dual Use
Education and Research are coupled!

Each dot is one tool
Size of dot indicates number of users

235 tools!

SUPREM
SPICE
Hubs ‘R Us

- Feb 2007: 1 hub
- Feb 2008: 5 hubs
- Feb 2009: 8 hubs
- Feb 2010: 21 hubs
- Sept 2010: >30 hubs
- Sept 2012: >40 hubs

Each hub has its own funding stream

Outside institutions:
EPA, NYSTAR, Rice
## 40+ Other Hubs

<table>
<thead>
<tr>
<th>Hub Name</th>
<th>Visitors</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>nanoHUB.org</td>
<td>234,949</td>
<td>481,906</td>
</tr>
<tr>
<td>NEEShubs.org</td>
<td>33,620</td>
<td>150,883</td>
</tr>
<tr>
<td>GlobalHUB.org</td>
<td>32,965</td>
<td>60,523</td>
</tr>
<tr>
<td>PharmaHUB.org</td>
<td>23,425</td>
<td>33,447</td>
</tr>
<tr>
<td>VHub.org</td>
<td>5,385</td>
<td>23,444</td>
</tr>
<tr>
<td>GCCHUB.org</td>
<td>3,391</td>
<td>14,667</td>
</tr>
<tr>
<td>iemHUB.org</td>
<td>1,224</td>
<td>12,550</td>
</tr>
<tr>
<td>C3Bio.org</td>
<td>1,863</td>
<td>10,995</td>
</tr>
<tr>
<td>HABRIcentral.org</td>
<td>1,173</td>
<td>10,490</td>
</tr>
<tr>
<td>StemEdHub.org</td>
<td>1,879</td>
<td>10,358</td>
</tr>
<tr>
<td>ciHUB.org</td>
<td>3,141</td>
<td>9,792</td>
</tr>
<tr>
<td>CleerHUB.org</td>
<td>1,635</td>
<td>9,666</td>
</tr>
</tbody>
</table>

~800,000 visitors total
Scientific Knowledge Transfer on nanoHUB.org
Making Research Useful for Others

Over 230,000 Users Annually

857 papers

14,000 students

HUBzero & Rappture

What’s Next?

Accessible (FREE, no installation)

In Research

In Education, Accelerate Innovation

Developer Friendly

User Friendly

260+ tools

>300 developers

12,000 Sim. Users Annually
Imagine

Simulation Tools

• Used by researchers
• Used by experimentalists
• Used in education

In a scientific cloud

Without any installation

Fully operational 24/7

Many proposals read alike

With assessed IMPACT

The PowerPoints are identical
Simulation Tools

- Used by researchers
- Used by experimentalists
- Used in education

In a scientific cloud
Without any installation
Fully operational 24/7

Many proposals read alike
With assessed IMPACT

The PowerPoints are identical

We achieved that dream.
Simulation Tools

• Used by researchers
• Used by experimentalists
• Used in education

In a scientific cloud

Without any installation

Fully operational 24/7

Many proposals read alike

With assessed IMPACT
Simulation Tools and Experimental Data

- Used by researchers
- Used by experimentalists
- Used in education

In a scientific cloud
Without any installation
Fully operational 24/7
With assessed IMPACT

Many proposals read alike
The PowerPoints are identical
Simulation Tools and Experimental Data

- Used by researchers
- Used by experimentalists
- Used in education

In a scientific cloud
Without any installation

Fully operational 24/7
In all areas of Nano Engineering and Science
With assessed IMPACT

Personalized Learning at all workforce levels

Become Part of the Day-to-Day Workflow

Many proposals read alike

The PowerPoints are identical
Simulation Tools and Experimental Data

- Used by researchers
- Used by experimentalists
- Used in education

In a scientific cloud
Without any installation
Fully operational 24/7
With assessed IMPACT

In all areas of Nano Engineering and Science
Personalized Learning at all workforce levels
Become Part of the Day-to-Day Workflow
Scientific Knowledge Transfer on nanoHUB.org
Making Research Useful for Others

Over 230,000 Users Annually
857 papers
14,000 students

Accessible (FREE, no installation)

In Research
In Education, Accelerate Innovation

Developer Friendly

User Friendly

HUBzero & Rappture

Success Criteria

260+ tools
>300 developers
12,000 Sim. Users Annually
Any Science Gateway’s Dream

Why is it so hard?
Any Science Gateway’s Dream
There are worlds between...

Research

Basic Research; Invention

“Valley of Death”

Applied Research; Innovation
7 Criteria for Successful Science Gateways
1: Outstanding Science

“Stuff the world wants”

Leveraged Research

$5.1M
2: Commitment to Dissemination

“faculty that want to give it away”

46 faculty

+ 6 site leads

106 grad students
Faculty Incentives

Tool Usage ≈ reading papers

17 tools
- 7,835 users
- 115 citations

Proof of Impact!
Great in Proposals!
Shaikh Ahmed

- Infused nanoHUB into existing classes
- Built a new nanoelectronics curriculum
- Used nanoHUB for research

Next Generation Faculty:

Post Doc at Purdue
Faculty at SIUC

- Infused nanoHUB into existing classes
- Built a new nanoelectronics curriculum
- Used nanoHUB for research

Early Tenure Promotion

6,183 users
8 tools
3: Technology for Dissemination

“simple and utterly dependable”
Less than 20 hours downtime last year!

Basic Research; Invention

$1M/year operation and bridge building

“Valley of Death”
Typical Dissemination Paths

- Customers
- Knowledge
  - Courses
  - Textbooks
  - IP
  - Publication
  - Seminar
  - Sim. Tools
    - Instruments
    - Data
  - Mod & Simulation
    - Experiments
    - Theory
  - Web Content

Problems:
- REALLY LONG stove pipe
- Web content: afterthought usually stale
- Data shared by email
- Tools spread by hiring
nanoHUB Technical Solution

Problems:
- **REALLY** LONG stove pipe
- Web content: afterthought usually stale
- Data shared by email
- Tools spread by hiring
4: Tech Transfer Processes

“dedicated technical site leads”

Consultants

Content Creation and Support

$2.2M

Significant portion of budget

“Valley of Death”
5: Understanding Stakeholders

Geek

Basic Research; Invention

CI Ops

nanoHUB.org

User

Applied Research; Innovation

“Valley of Death”
5: Understanding Stakeholders

- Geek
- User
- CI
- Ops

Basic Research; Invention

Virtual Org.

“Valley of Death”

Applied Research; Innovation
6: Open Assessment / Incentives

“gather, understand, disseminate stats”

Access, Use, Impact
7: Business Model

Research: Sustained Academic Funding

Product Development: Real Business Plans

Basic Research; Invention

"Valley of Death"

Applied Research; Innovation