# DeConstructing the Massachusetts Green High Performance Computing Center in Holyoke

Jim Kurose
Department of Computer Science
Exec. Assoc. Dean, College of Natural Sciences
Faculty Advisor for Science and Engineering, VCRE
University of Massachusetts Amherst

DeConstructing the Massachusetts Green High Performance Computing Center in Holyoke

*What is HPC?*

*What might this mean for Holyoke, the Pioneer valley, and the Commonwealth?*

*What does it mean to be "green?"*

*What does an HPC data/computing center look like?*

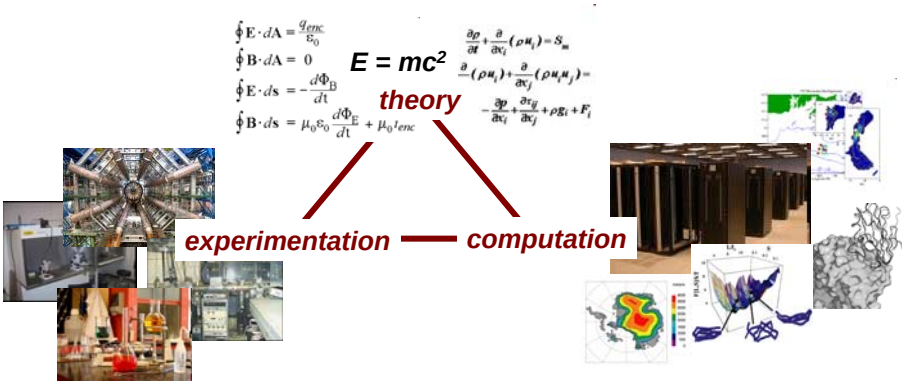
# Overview*

- What is HPC?
- What does an HPC center look like?
- What does it mean to be green?
- What might this mean for Holyoke, The Pioneer Valley, and the Commonwealth?

**Disclaimer:* The views expressed in the presentation are purely those of the speaker, and thus do not necessarily reflect the views of his employer, colleagues, friends, family, the GHPCC, the Commonwealth of Massachusetts, or indeed anyone else……

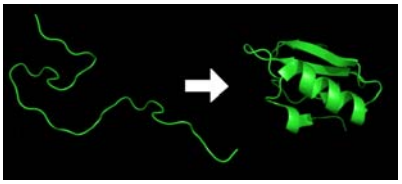
# What is HPC?

- using large clusters of computers, or special-purpose "supercomputers" in computing applications that are too large or would take too long for "traditional" computing
- "third leg" of science, engineering inquiry

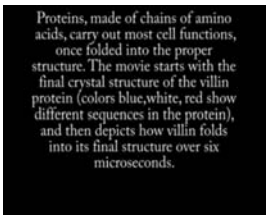


# HPC application: protein folding

a protein begins as a long sequence of amino acids…..



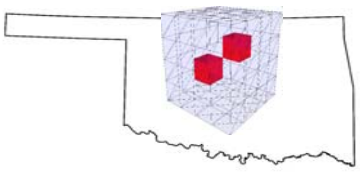
…. which then ***folds*** into a characteristic 3D functional structure for that protein

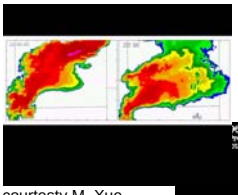



- understanding protein folding key to understanding:
  - *disease:* mis-folding linked to Alzheimer's, mad cow, others..
  - *drug design:* protein flexibility
- protein folding takes microseconds in vivo, but 100's CPU days to simulate: HPC

# HPC application: numerical weather prediction

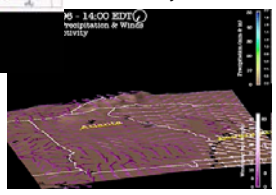
- computer simulation of weather/atmosphere evolution over time
- divide atmosphere into 3D grid, detailed simulation within each small (e.g., $10m^3$), gridpoint, couple with neighbors
- CASA: collaborative adaptive sensing of the atmosphere: UMass-led NSF Engineering Research Center



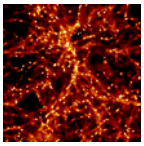


courtesy M. Xue

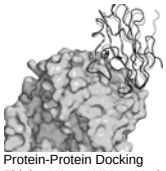
courtesy L. Trenish



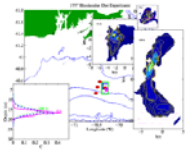
# HPC: Sampling UMass Activities



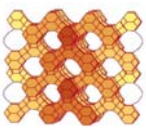
Cosmology simulations,
N. Katz, UMass/Amherst



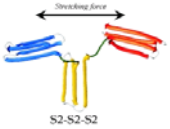
Protein-Protein Docking
Zhiping Weng, UMassMed



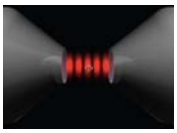
Ocean Mixing, Miles Sundermeyer
UMass/Dartmouth



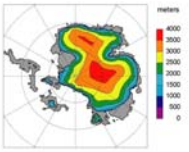
Proton Hopping in Organic Molecules
S. Auerbach, UMassAmherst




Protein Unfolding, Valeri
Barsegov, UMass/Lowell



Cooling a Single Atom in an Optical
Cavity, Kurt Jacobs, UMass/Boston

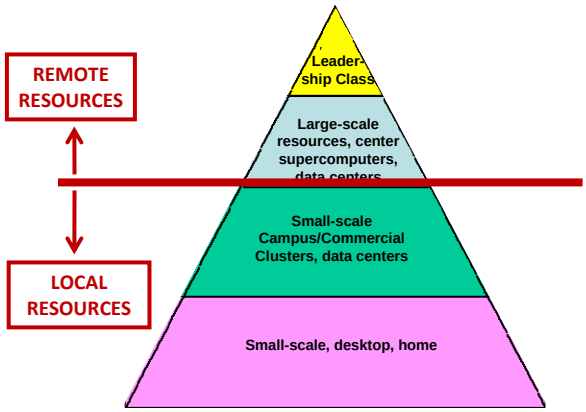


Climate change models, R. Decanto
UMass/Amherst

# Many (many!) other HPC application areas

- ***energy:*** fuel cells, underground oil/gas location, flow
- ***materials:*** high-temperature superconductors, efficient catalysts
- ***astronomy:*** origins of the universe
- ***physics:*** low-temperature physics, high-energy physics (LHC)
- ***finance:*** financial modeling, trading analytics
- ***environmental modeling:*** earth, atmosphere, water
- ***industry product design:*** modeling versus physical prototyping, optimization
- ***industry services***: complex data-driven decision-making, risk analysis, supply chain optimization, optimized routing

## Branscomb Pyramid: classes of HPC facilities




"most science is conducted at lower levels of the pyramid, using desktop computers, laboratory clusters and university-scale computing infrastructure…." [Dan Reed Microsoft Corporate VP for Technology Policy and Strategy]

## Overview

- What is HPC?
- What does an HPC center look like?
- What does it mean to be green?
- What might this mean for Holyoke, The Pioneer Valley, and the Commonwealth?

# Data Centers: from large……

- Microsoft (Chicago):
  - 170,000 sq ft
  - 56 *container* slots, up to 2K servers per container
  - currently: 30 MW
  - inexpensive electricity, Gbit Xroads

Google:

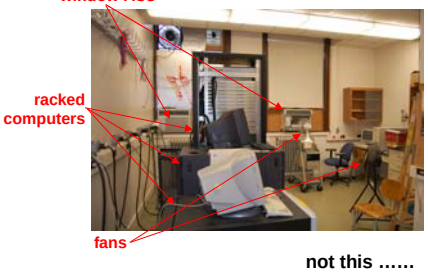
- not much public info available
- video tour, Google's first container-based data center, 2009:

http://www.youtube.com/watch?v=zRwPSFpLX8I

Microsoft Chicago Data Center: containers 1st floor, traditional server rooms, 2nd floor (cnet.com)

Google data center, Columbia River, The Dalles, Ore (source: IEEE Spectrum)

# Data Centers: …. to the not so big




**not this ……**

**…. this**

- UMass HPC co-location center
- 23 water-cooled racks
- <short video walkthrough>

## Data Centers: choosing a location

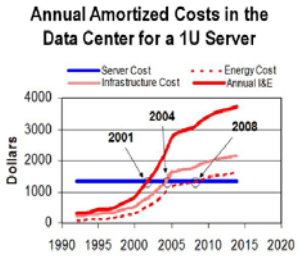
- *Google:* "The Columbia Gorge area has the right combination of *energy infrastructure*, *developable land* and *available workforce* for the data center. Additionally, the *team from the local community* that welcomed us has worked exceptionally hard throughout the selection process. All of this has truly distinguished The Dalles as a great place for our data center."
- also:
  - network connectivity
  - traditional considerations: housing, cost of business, schools, higher ed institutions,….

## Overview

- What is HPC?
- What does an HPC center look like?
- What does it mean to be ***green***?
- What might this mean for Holyoke, The Pioneer Valley, and the Commonwealth?

## Green Computing: reducing IT's energy footprint

- computing, communication infrastructure (data centers, telecom, personal devices):
  - 3% total carbon emissions; 6% annual growth, rising to 1.43 Gt $CO_2e$ by 2020
  - data center power, cooling: expensive, large energy consumption




*goal:* reduce energy footprint of computing infrastructure

## Reducing IT's energy footprint: source

- green/clean electricity sources for data centers lower carbon footprint
  - hydro: renewable source, cooling (geothermal also)
  - solar, wind, biomass: additional sources, lowering carbon footprint
- IT industry move towards cheap, green energy


*Google data center, Columbia River, The Dalles, Ore (source: IEEE Spectrum)*


*Holyoke Dam (Holyoke Gas & Electric, Conn. River)*

## Reducing IT's energy footprint: data center

green technologies for data center infrastructure:

- smart / green building technologies
- smart air-conditioning
  - selective cooling data center
  - intelligent air-flow cuts power costs for cooling
- smart design: smart server racks for better heat dissipation

Microsoft genomote, temperature variation across server racks

Studies show can save $1-2M/year **per typical/modest (100-rack) data center**

There are 1000's of small to large data centers in MA

## Reducing IT's Energy footprint: energy-efficient computing

- design energy-efficient *chips, servers, systems*
  - smart chips that consume less power
  - smart servers that use software to run slower or turn off when idling
  - virtualization technologies: move server load
    - from servers running hot to servers running cool
    - to minimize number of operating servers
- industry, academia in Massachusetts are working on next-generation of energy-efficient computing technologies

# HPC center: energy savings

- Power Usage Efficiency (PUE): data center efficiency = $\frac{\text{Power entering HPC facility}}{\text{Power delivered to computers}}$

- 2006: avg. data center: PUE = 2 *[EPA]*
  - state-of-the-art green technologies can reduce 2011 PUE to 1.2 *[EPA]*
  - smart cooling technologies: $1-2 M annual savings in 30K sq-ft data center *[HP ]*
  - green data centers: recover initial additional investment costs in 2-3 years *[IBM]*

| Scenario | PUE |
|---|---|
| Current Trends | 1.9 |
| Improved Operations | 1.7 |
| Best Practices | 1.3 |
| State-of-the-Art | 1.2 |

*Figure 1: EPA Estimated PUE Values in 2011* [5]

# Green Computing: enabling savings in other sectors

smart technologies:
- ind. automation (0.97 $GtCO_2e$ savings)
- smart logistics (1.52 $GtCO_2e$ savings)
- smart buildings (1.68 $GtCO_2e$ savings)
- smart grids (2.03 $GtCO_2e$ savings)

"While the [IT] sector plans to significantly step up the energy efficiency of its products and services, ICT's largest influence will be by enabling energy efficiencies in other sectors, an opportunity that could deliver carbon savings five times larger than the total emissions from the entire ICT sector in 2020." (www.smart2020.org)

**SMART 2020: Enabling the Low Carbon Economy in the Information Age**

# What is the "G" in GHPCC?

## Green energy sources

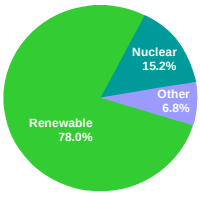



Figure 1: HG&E 2008 Energy Sources (2008)

## Green facility operations

- state-of-the-art efficiency
- federated computing

## Green computational science and engineering

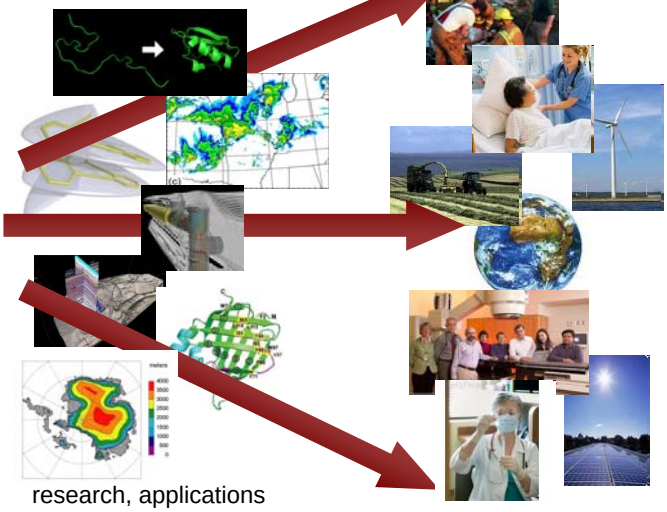
- energy sciences
- environmental sciences
- green computing

# Overview

- What is HPC?
- What does an HPC center look like?
- What does it mean to be green?
- What might this mean for Holyoke, The Pioneer Valley, and the Commonwealth?

# What is a GHPCC?

computing



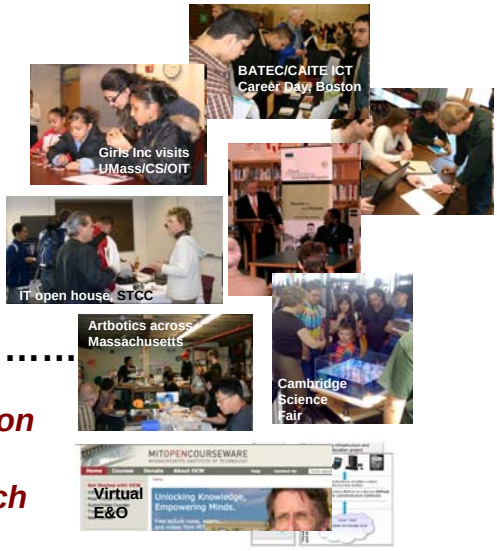
research, applications

people

# What is a GHPCC?

**Not just this……**

**this……**




***Education and Outreach***


## Who was there?

HCC (9), UMass Amherst (5), Mount Holyoke College (4), Boston University (3), EMC Corporation (3), Holyoke Public Schools (3), MIT (3), Northeastern University (3), Others (3), ISO New England (3), STCC (2), UMass System (2), ICT Center, STCC (2), Nuestras Raices (2), Art of Politics, Boys & Girls Club of Greater Holyoke, Inc., Community Education Project, Congressman John Olver's Office, Dean Technical High School, Five Colleges, Inc, Girls Incorproated of Holyoke, Hampden WIB, HCC - Tech Prep, Holyoke Works, Massachusetts Broadband Institute, Mt.. Auburn Associates, MTC's John Adams Innovation Institute, Open Square Properties LLC, Peck School, Renaissance Computing Institute, Rensselaer Polytechnic Institute, UMass Medical School, UMass Boston, UMass Donahue Institute, UMass Lowell


## What did we hear?

**Breakouts:** stakeholders were represented in each group

- K12, community and economic development organizations: Bill Ennen (MTC) and Carol Soules (UMass)
- Community college partners: Sue Mackler (HCC) and Gordon Snyder (STCC/ICT Center)
- University partners: Chris Hill (MIT) and Larry Finkelstein (NU)


## K-12, community

- access in Canal District : connectivity, learning about computing, community meetings (cross-cultural, multi-generational)
- space, staff for outreach
- education
  - professional development (teachers)
  - family involvement
  - college access/awareness
  - adult education


## Community Colleges

- tours, professional development, “Space Camp”-like experience, student mentoring
- boot camps - awarded certificates, involving high level GHPCC faculty
- identify currently available courses/curriculum/content/programs
- new programs (e.g., green computing). ATE center?
- student, faculty access to new equipment/technology
- access to subject matter expertise e.g. Storage, Networking including NAS and SAN, Virtualization, Elastic Computing, Data, etc
- certification programs (with EMC, Cisco)

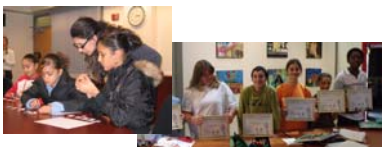

## Universities

- "science in the cloud": on demand, virtual resource in K-12, community colleges
- enhancing student pipeline
- connecting community colleges, K12 schools to university research

## C4 - Canals & Computers, Community & Collaboration

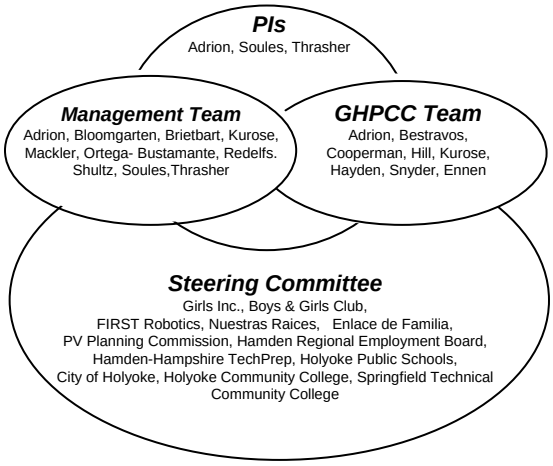
- NSF CISE BPC proposal that builds on GHPCC, Five-College Community-Based Learning, the "5+2 colleges," CAITE

- direct involvement of HPS, CBOs
- education
  - college readiness, career awareness
  - Science in the Cloud



- outreach
  - access, events, Mminigrants
- training
  - vocational, ABE and Community College certification
- pathways

# C4 - Canals & Computers, Community & Collaboration




Direct community input, strong leadership team

# GHPCC: Catalyst for Economic Development

- GHPCC not, by itself, solution for creating regional economic growth
- part of larger strategy for economic development in city/region, especially Innovation District
  - anchor" for Innovation District
  - catalyst and "proof point" to private industry of Holyoke as location energy-intensive IT facilities
- commitment by 4 leading research universities, partners, state to put Holyoke "on the radar screen" for IT companies, others

## DeConstructing the Massachusetts Green High Performance Computing Center in Holyoke

- HPC as enabler of science, engineering
- HPC centers: computational infrastructure for 21st century
- Green HPC: leveraging, saving resources
- Opportunities for Holyoke, the Pioneer Valley, and the Commonwealth

## Thank you!

?? || /* */

# UMass Amherst: state leadership in IT education

**CITI:** Commonwealth IT Initiative

- strengthening, modernizing, expanding IT education in *all segments* of public higher ed
  - building IT fluent workforce
  - culture of collaboration
- funded since 2000: BHE, foundations, industry, state
- programmatics:
  - IT across the curriculum
  - faculty development
  - curriculum enhancement

**FOR THE GOOD** Betty Medina Lichtenstein explains her needs to a University of Massachusetts class; from left, Jason De Furia, Prof. Carol Soules, Ryan Roy, Mankin Yuen and Erica Sawyer, a graduate assistant. The New York Times 12/29/09

Betty Medina Lichtenstein of Enlace de Familias, Holyoke, working with UMass students/faculty on IT capstone course

# UMass Amherst: state leadership in IT education

***CAITE:** Commonwealth Alliance for IT Education*

*Addressing under-representation in IT education and workforce, focusing on women and minorities in groups that are under-represented in the MA innovation economy.*

- funded by $4.1M NSF Broadening Participation in Computing grants
- focus on community colleges
- *outreach:* college/career fairs, campus visits, computing camps, IT internships
- *pathways:* intro courses, transfer agreements, peer mentoring, advising
- led by Rick Adrion, UMass/CS


caite.cs.umass.edu

## Possible education and outreach activities

- hands-on access to state-of-the-art, next-gen computing/networking systems from industrial partners (e.g., for academies, students/faculty training)
- industry advisory board for educational activities – education needs for next-gen products and services
- development fund pool for educational innovation in green HPC
- internships, externships with industrial partners
- pursuing federal funding
  - NSF ATE center planning grant
  - NSF CISE progam: Cyberinfrastructure Training, Education, Advancement, and Mentoring for Our 21st Century Workforce

## GHPCC: Catalyst for Economic Development

- GHPCC not, by itself, solution for creating regional economic growth
- Part of larger strategy for economic development in city/region, especially Innovation District
  - anchor" for Innovation District
  - catalyst and "proof point" to private industry that Holyoke is a place where they can develop energy-intensive IT facilities
- commitment by 4 leading research universities, partners, state to put Holyoke "on the radar screen" for IT companies, others