

The Center for Computational Research

Russ Miller

Director, Center for Computational Research

UB Distinguished Professor, Computer Science & Engineering

Senior Research Scientist, Hauptman-Woodward Medical Inst

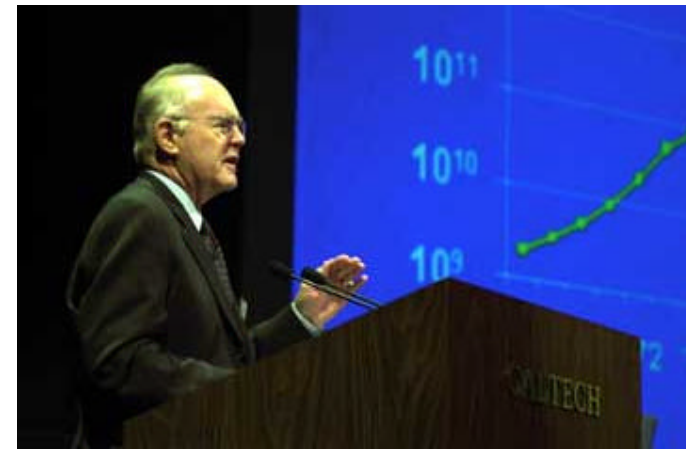


University at Buffalo

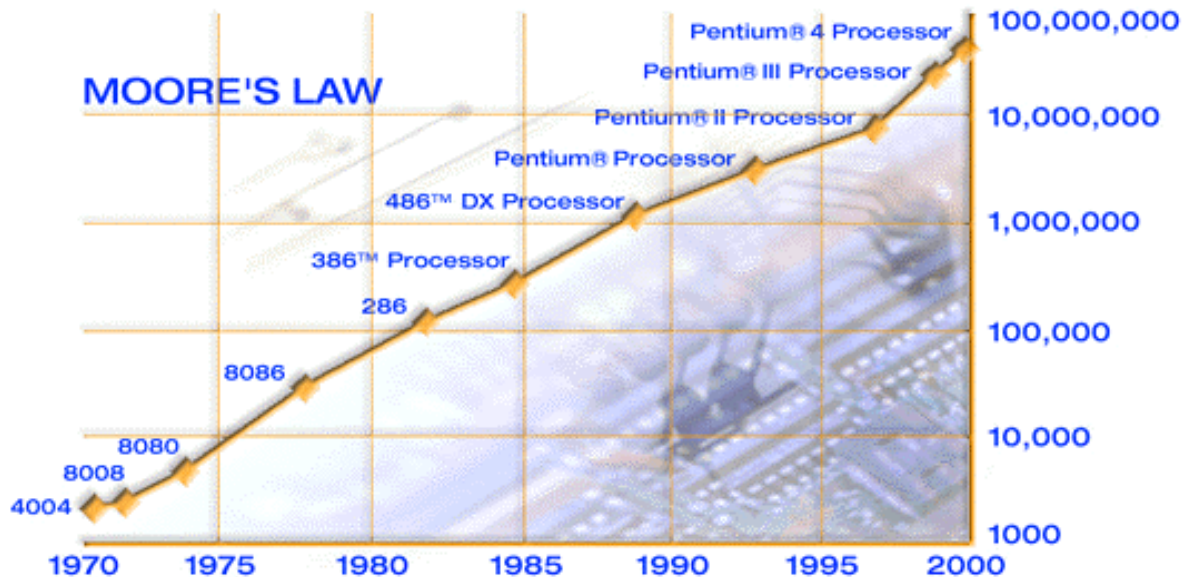
The State University of New York

Gordon E. Moore

- Co-Founder of Intel
- Predicted (1965/75) that transistor density would double every 12/18 months
- Processing speed doubling every 18 mos.
- Disk storage doubling every 12 mos.
- Aggregate bandwidth doubling every 9 mos.



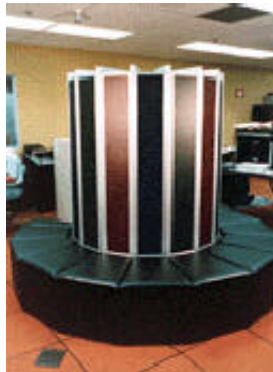
Gordon E. Moore



- A computation that took 1 year to run on a PC in 1985 would only take 5 mins to run on a PC today!
- A computation that runs in 2 hours on a PC today would have taken 24 years to run on a PC in 1985!

Supercomputers

- Fastest computers at any point in time
- Used to solve large and complex problems
- Machines 1000 times faster than a PC
- Machines 10 times slower than what you need to solve the most challenging problems



Cray1 - 1976



“Seymour Cray is the Thomas Edison of the supercomputing industry” **Seymour Cray**
- Larry L. Smarr 1925-1996

Beowulf Clusters

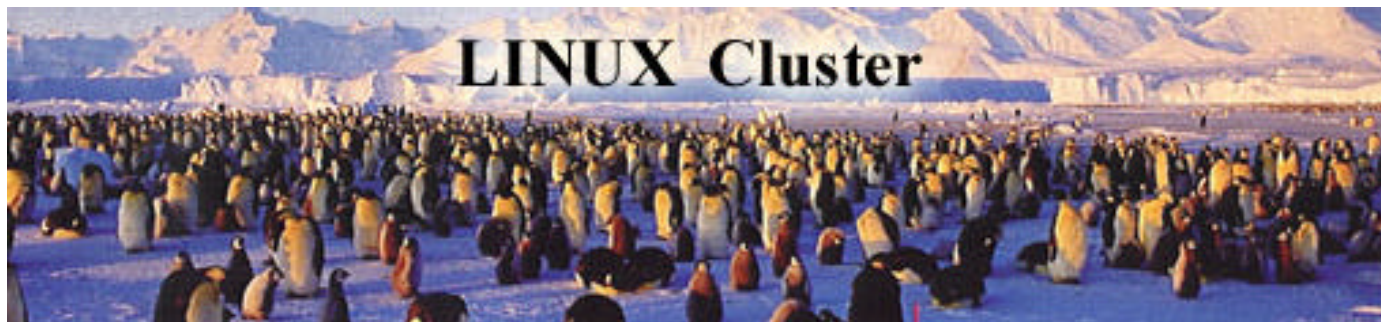
■ Industry Standard Hardware and Software

Thomas Sterling
Caltech

- ❑ PC-Based Components (Intel or AMD)
- ❑ Ethernet or Myrinet
- ❑ Linux, PBS, MPI
- ❑ “Commodity Off-The-Shelf” (COTS)

■ Operates as a Single System

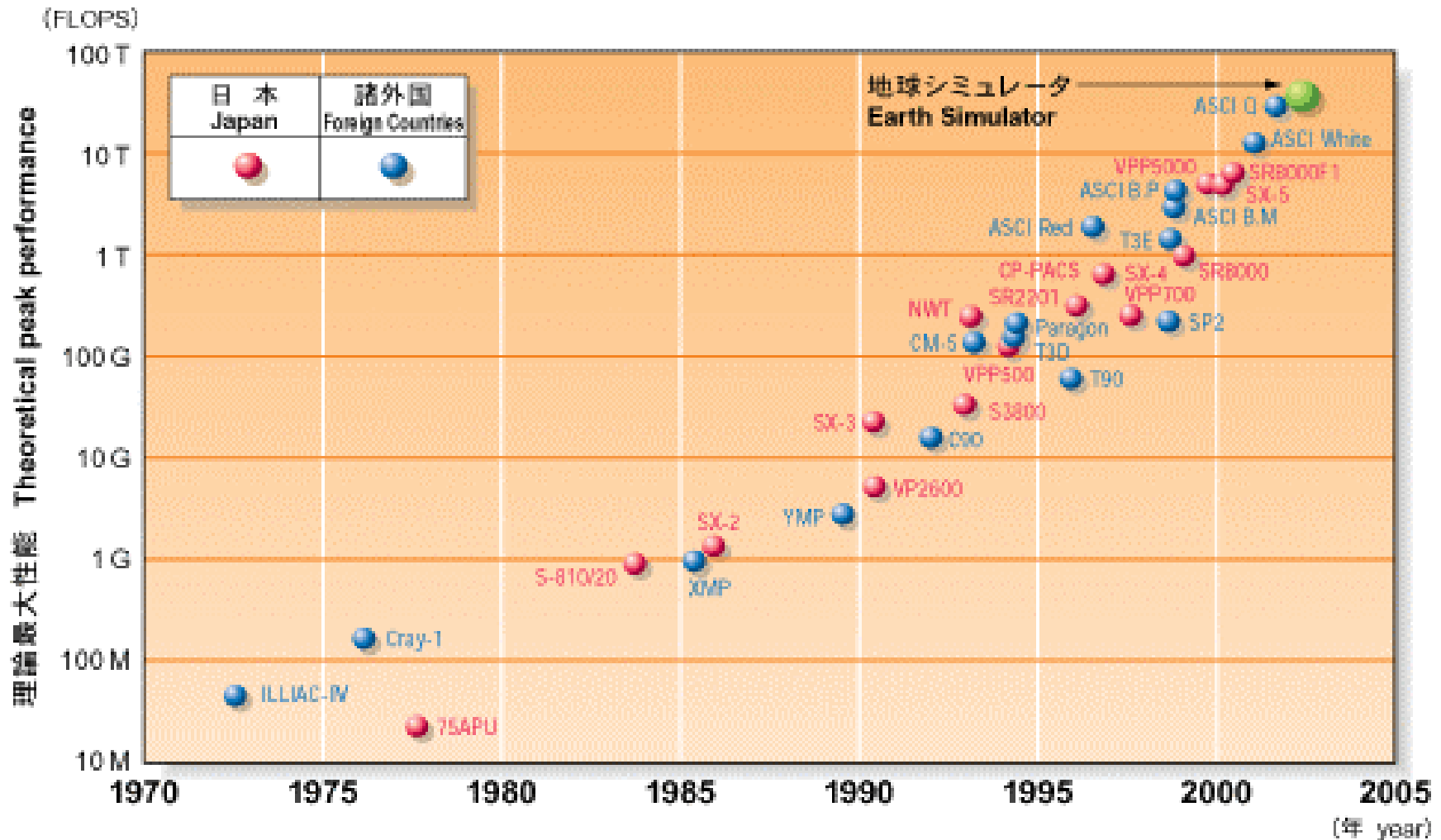
■ Rivals Performance of Traditional Supercomputer at a Fraction of the Price



Fastest Computers

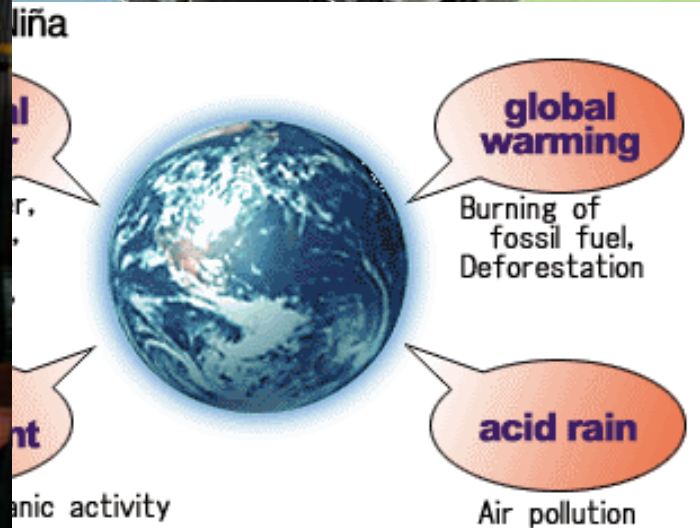
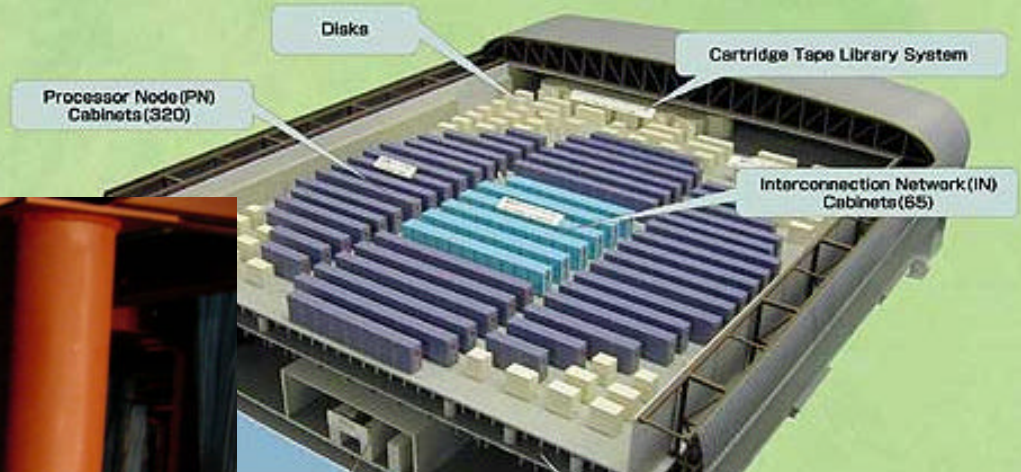
Year	Mach	Procs	GFlops		Year	Mach	Procs	GFlops
1976	Cray 1	1	0.1		1993	Cray T3D	1024	152
1982	Cray X-MP	4	0.9		1994	Fujitsu VPP	140	236
1986	Cray 2	4	2		1996	Hitachi SR2	2048	368
1989	Cray Y-MP	8	2.7		1997	Intel ASCI-R	9152	1830
1989	TMC CM-2	8192	28		1999	SGI ASCI-BM	6144	3072
1992	TMC CM-5	1024	131		2000	IBM ASCI-W	8192	12,288
<p>A 1-year calc in 1980 = 5.4 sec today A 1990 HPC = a laptop today</p>					2002	NEC E.S.	5120	40,960

Growth of Peak Performance



Earth Simulator

- 40TFlops Peak
- Homogeneous, Centralized,



Center for Computational Research 1999-2003 Snapshot

■ High-Performance Computing and High-End Visualization

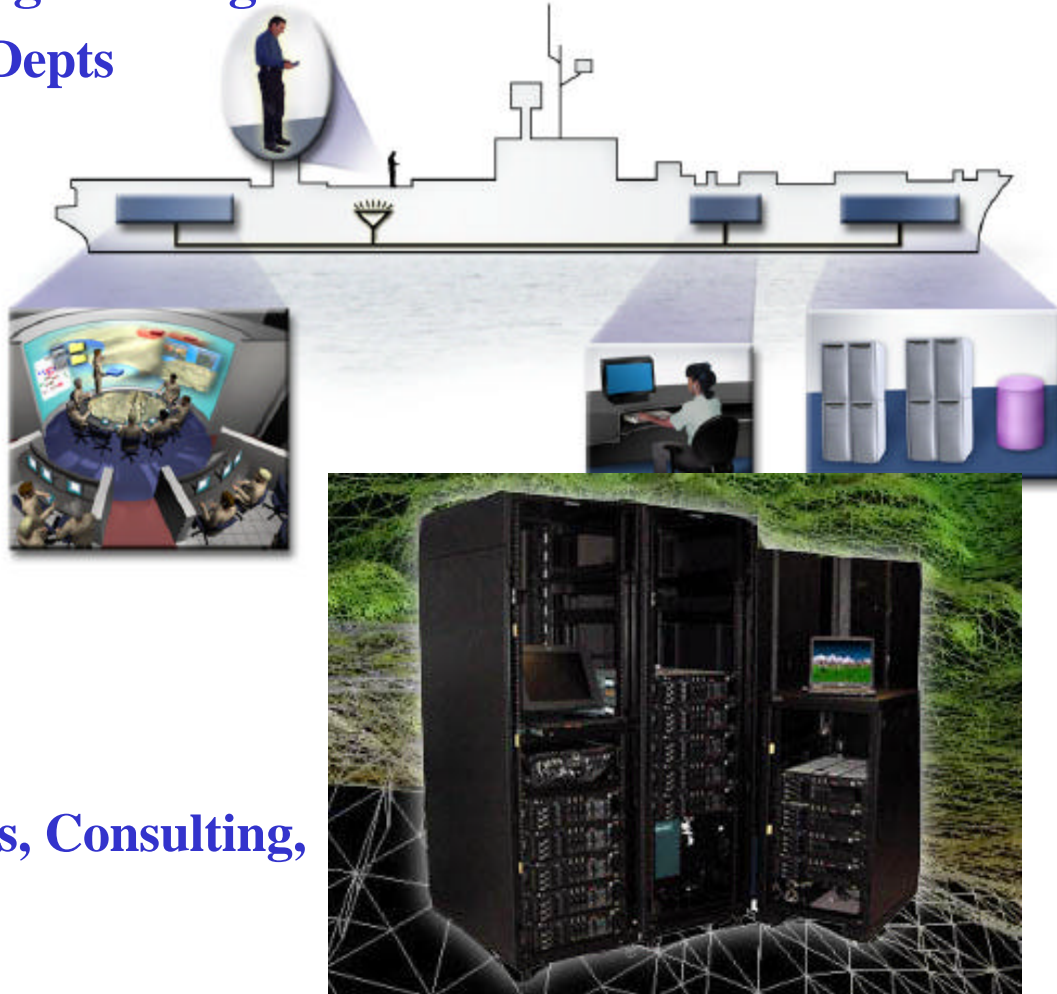
- ❑ 110 Research Groups in 27 Depts
- ❑ 13 Local Companies
- ❑ 10 Local Institutions

■ External Funding

- ❑ \$111M External Funding
 - \$13.5M as lead
 - \$97.5M in support
- ❑ \$41.8M Vendor Donations

■ Deliverables

- ❑ 350+ Publications
- ❑ Software, Media, Algorithms, Consulting, Training, CPU Cycles...



Major CCR Resources

- **Dell Linux Cluster: #22 ® #25 ® #38**
 - ❑ 600 P4 Processors (2.4 GHz)
 - ❑ 600 GB RAM; 40 TB Disk; Myrinet
- **Dell Linux Cluster: #187 ® #368 ® off**
 - ❑ 4036 Processors (PIII 1.2 GHz)
 - ❑ 2TB RAM; 160TB Disk; 16TB SN
 - ❑ Restricted Use (Skolnick)



■ SGI Origin3800

- ❑ 64 Processors (400 MHz)
- ❑ 32 GB RAM; 400 GB Disk

■ IBM RS/6000 SP

- ❑ 78 Processors
- ❑ 26 GB RAM; 640 GB Disk

■ Sun Microsystems Cluster

- ❑ 48 Sun Ultra 5s (333MHz)
- ❑ 16 Dual Sunblades (750MHz)
- ❑ 30 GB RAM, Myrinet



■ SGI Intel Linux Cluster

- ❑ 150 PIII Processors (1 GHz)
- ❑ 75 GB RAM, 2.5 TB Disk Storage



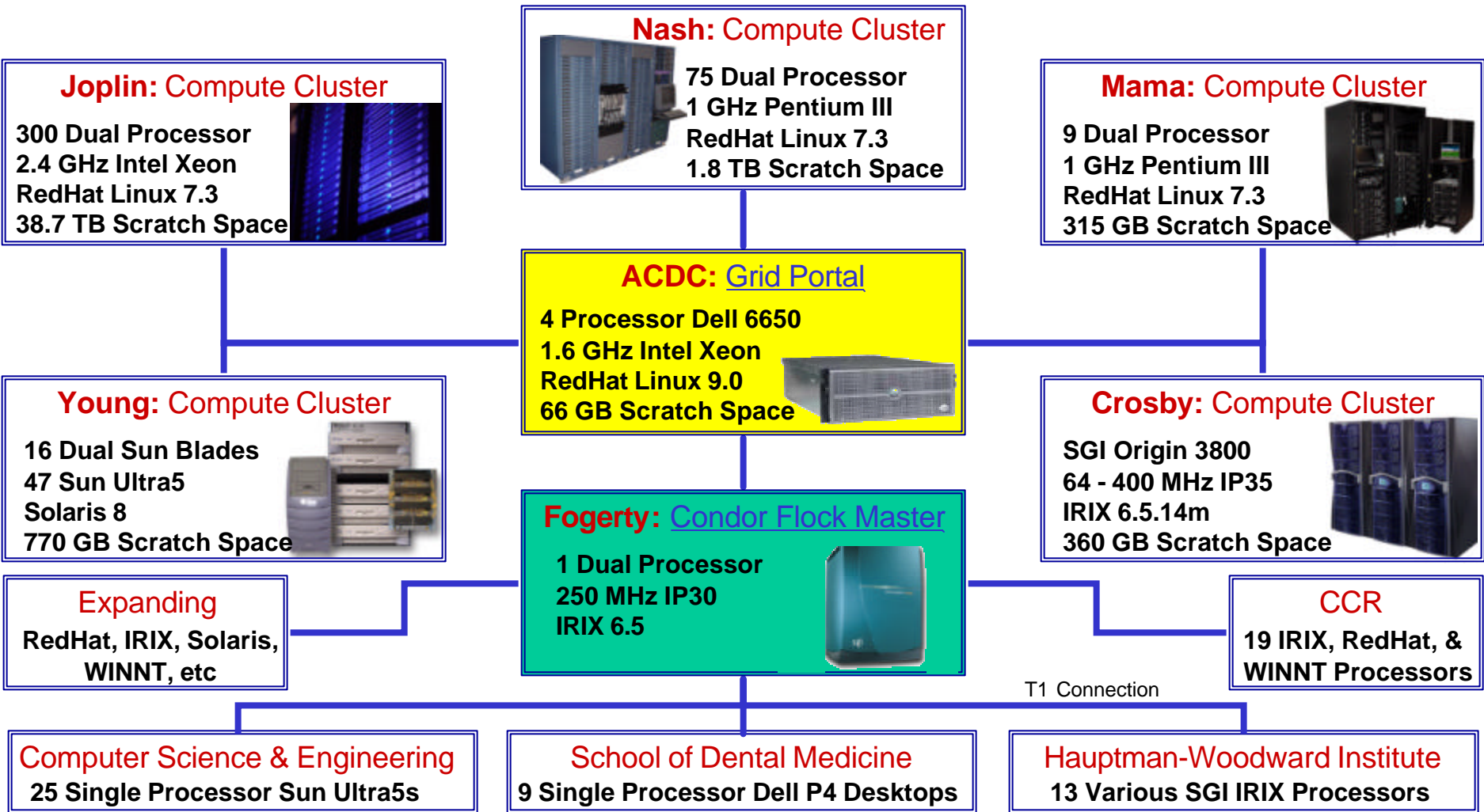
■ Apex Bioinformatics System

- ❑ Sun V880 (3), 6800, 280R (2), PIIIs
- ❑ Sun 3960: 7 TB Disk Storage

■ HP/Compaq SAN (4Q03)

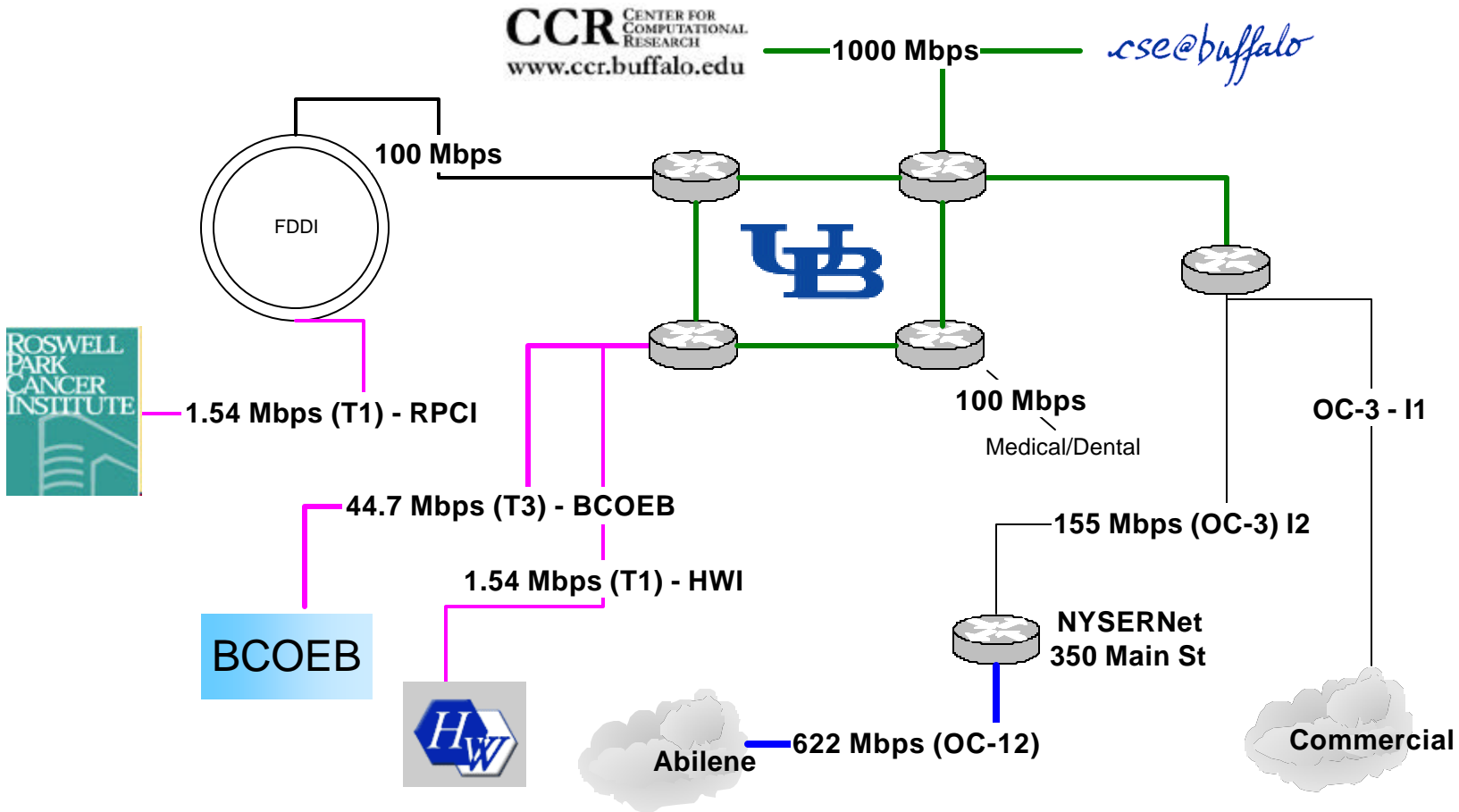
- ❑ 75 TB Disk; 200 TB Tape

Advanced CCR Data Center (ACDC) Computational Grid Overview



Note: Network connections are 100 Mbps unless otherwise noted.

Network Connections

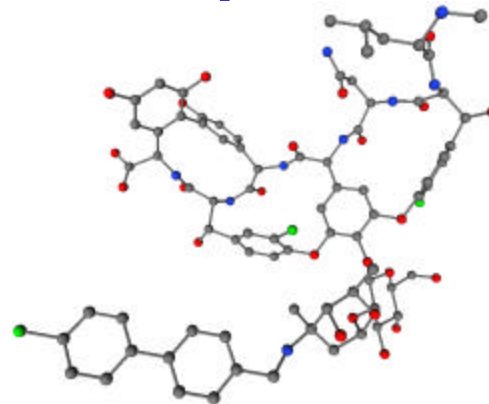
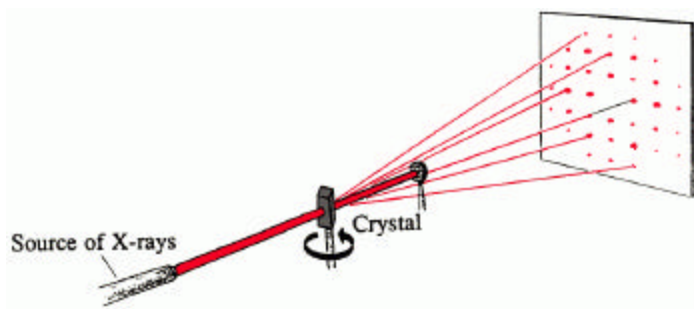
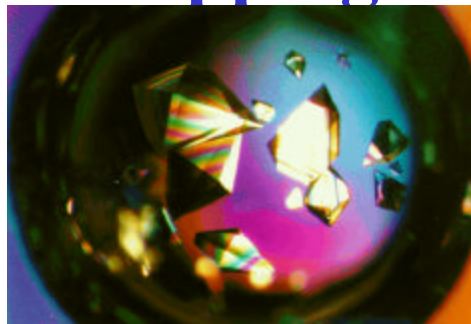


X-Ray Crystallography

- **Objective: Provide a 3-D mapping of the atoms in a crystal.**

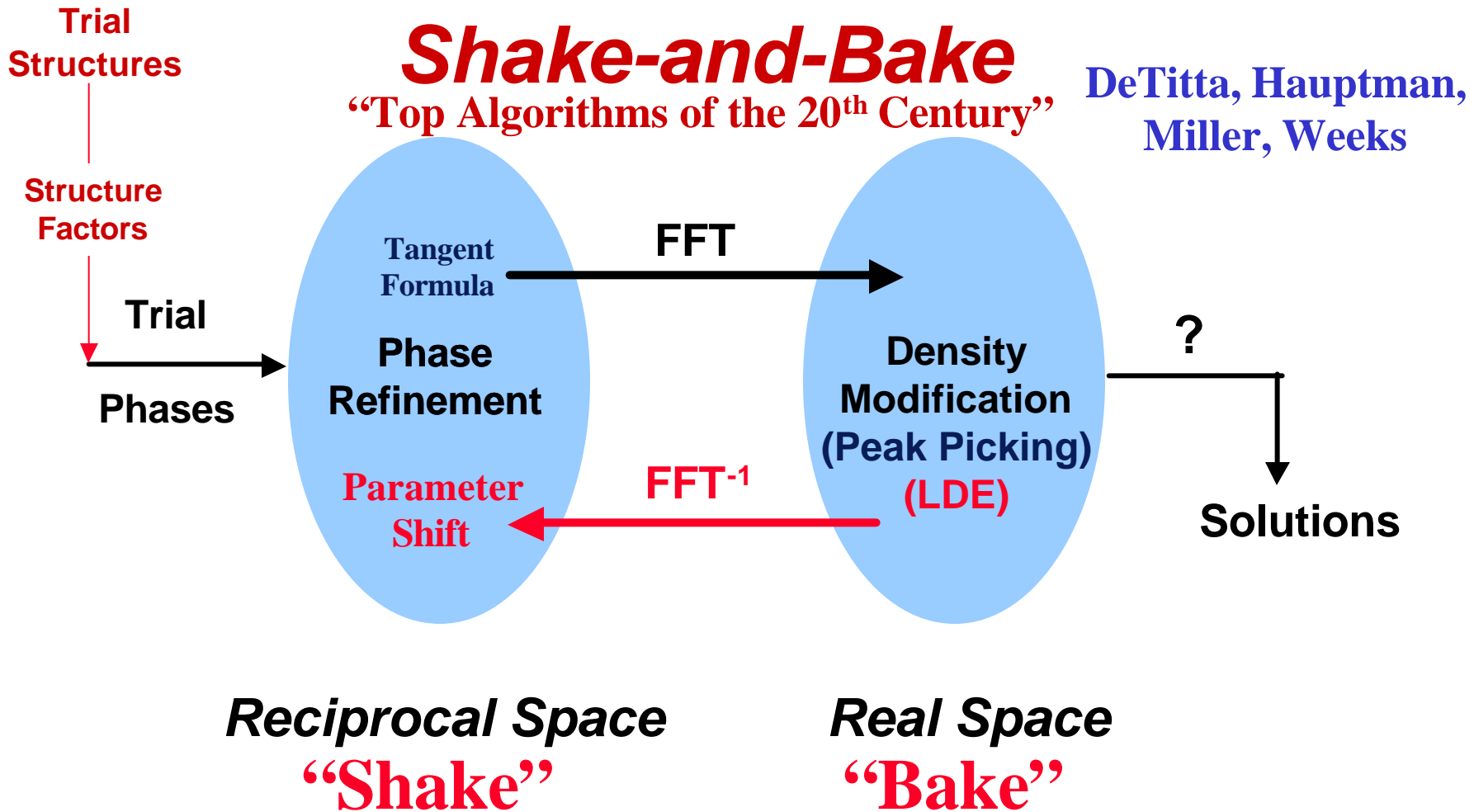
- **Procedure:**

1. **Isolate a single crystal.**
2. **Perform the X-Ray diffraction experiment.**



3. **Determine molecular structure that agrees with diffraction data.**

Shake-and-Bake Method: Dual-Space Refinement



Phasing and Structure Size

Se-Met with *Shake-and-Bake*?

Se-Met

190kDa

Multiple Isomorphous Replacement?

Shake-and-Bake

Conventional Direct Methods

Vancomycin



Number of Atoms in Structure

Grid-Based *SnB* Objectives

- **Install Grid-Enabled Version of *SnB***
- **Job Submission and Monitoring over Internet**
- ***SnB* Output Stored in Database**
- ***SnB* Output Mined through Internet-Based Integrated Querying Tool**

- **Serve as Template for Chem-Grid & Bio-Grid**
- **Experience with Globus and Related Tools**

Grid Services and Applications

**ACDC-Grid
Computational
Resources**

Applications

Shake-and-Bake

Apache

MySQL

Oracle

High-level Services and Tools

Globus
Toolkit

NWS

MPI

MPI-IO

C, C++, Fortran, PHP

globusrun

Core Services

Metacomputing
Directory
Service

Globus
Security
Interface

GRAM

GASS

Local Services

Condor

Stork

MPI

RedHat Linux

WINNT

LSF

PBS

Maui Scheduler

TCP

UDP

Irix

Solaris

**ACDC-Grid
Data
Resources**

Adapted from Ian Foster and Carl Kesselman



ACDC-Grid Portal

CCR Grid Computing Services: - Microsoft Internet Explorer

File Edit View Favorites Tools Help

UB University at Buffalo The State University of New York

CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

PORTAL LOGIN

- Grid General Info
 - » About ACDC Grid
 - » Computational Grid
 - » Data Grid
 - » Publications
 - » Technical Papers
 - » Presentations
 - » Contact Us
 - » Grid Account Request
 - » Grid Account Support
 - » Events
 - » News
- Projects
- Resources
- Education/Outreach
- Staff Only
- CCR HOME

Welcome to Grid Computing Services

University at Buffalo Center for Computational Research is currently forming the first Western New York computational grid. The computational grid consist of many supercomputers located at the Center and several other networked supercomputers throughout the Western New York region. These resources will be shared by many researchers from several departments working on a diverse suite of problems including Bioinformatics, Computational Chemistry, and Medical Imaging to name a few.

We also provide grid computing support for the University's Center for Computational Research learning & teaching and research activities plus the infrastructure for both high performance computing and grid enabled software.

Got your "Grid Computing Guide"?

Do you want to learn about 'Grid Computing'?

Advanced
Center for Computational Research
Data
Center



Data Grid Capabilities

The screenshot shows a web browser window titled "CCR Grid Computing Services: Data Management - Microsoft Internet Explorer". The page header includes the University at Buffalo logo and the text "Center for Computational Research GRID PORTAL High Performance Grid Computing".

On the left side, there is a "PORTAL LOGOUT" menu with categories: "User Tools" (Manage Account), "Grid General Info", "Projects", "Resources" (Computational Grid, Job Submission, Job/Queue Status, Data Grid, Network Status, Running/Queued Jobs, PBS Job History, Grid Portal Statistics, Condor Flock Statistics, User Information), "Education/Outreach", "Staff Only", and "CCR HOME".

The main content area features a file browser interface. At the top, there are dropdown menus for "VIEW" (set to "Group"), "GROUP" (set to "miller"), and "UserList" (set to "rappleye"). Below these, a tree view shows the following structure:

- rappleye
 - KeyMaster
 - Morpheus
 - Tank
 - Agent
 - Rabbit
 - Tank
 - Morpheus
 - Oracle.m
 - Neo
 - Cypher
 - Neo
 - Morpheus
 - Oracle

A yellow callout bubble points to the "Oracle.m" file, containing the text: "Browser view of 'miller' group files published by user 'rappleye'".

At the bottom of the page, there is a decorative footer with the text "Advanced Center for Computational Research Data" and a stylized "GRID" logo.

Grid Portal Job Status

■ Grid-enabled jobs can be monitored using the Grid Portal web interface dynamically.

□ Charts are based on:

- total CPU hours, or
- total jobs, or
- total runtime.

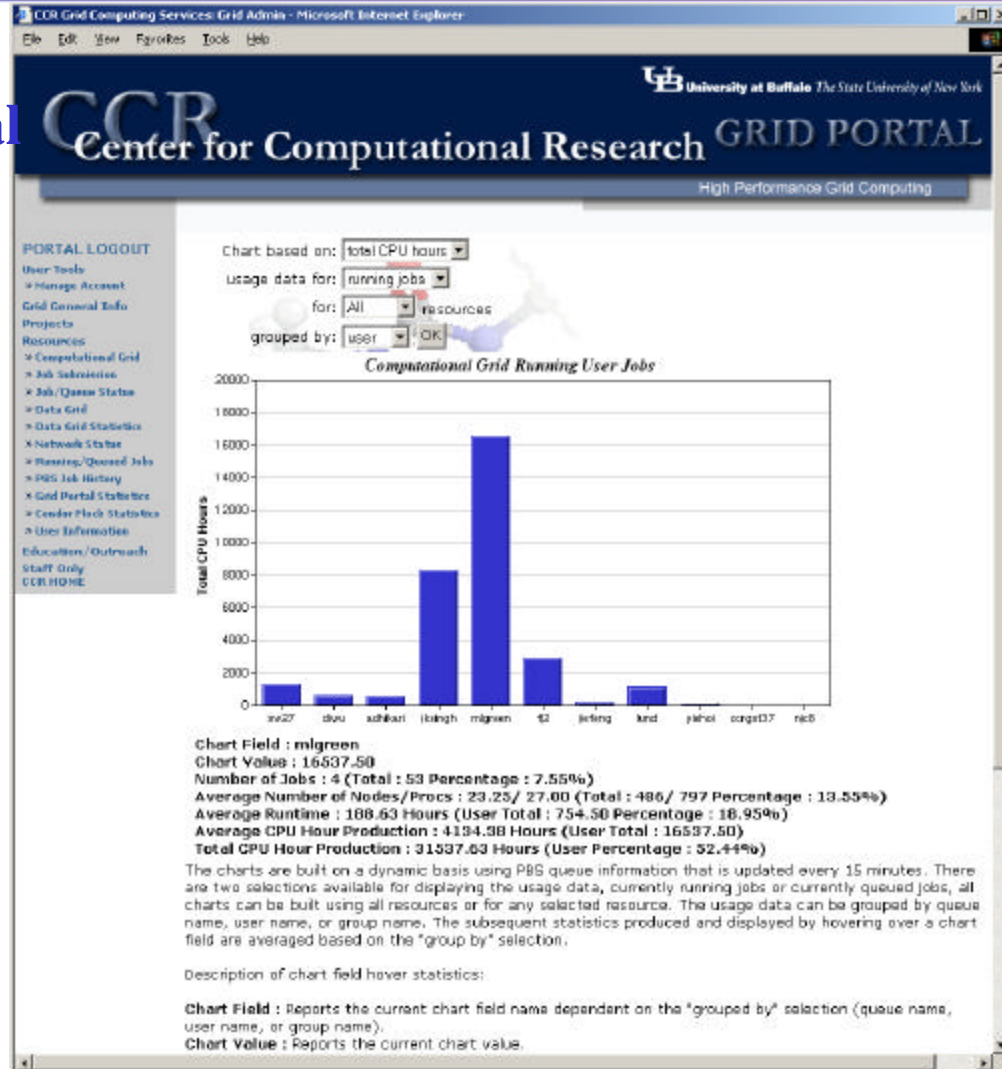
□ Usage data for:

- running jobs, or
- queued jobs.

□ Individual or all resources.

□ Grouped by:

- group, or
- user, or
- queue.



Grid Portal Job Status

CCR Grid Computing Services: Grid Admin - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CCR University at Buffalo The State University of New York
Center for Computational Research GRID PORTAL
High Performance Grid Computing

PORTAL LOGOUT
User Tools
Manage Account
Grid General Info
Projects
Resources
Computational Grid
Job Submissions
Job/Queue Status
Data Grid
Data Grid Statistics
Network Status
Running/Queued Jobs
PBS Job History
Grid Portal Statistics
Center Fleet Statistics
User Information
Education/Outreach
Staff Only
CCR 8048

Chart based on: total CPU hours
usage data for: queued jobs
for: All resources
grouped by: user

Computational Grid Queued User Jobs

User	Total CPU Hours
mlgreen	7623.17
ksingh	10800
sa27	500
dsingh	6500
ritdici	200
chen938	3000
janovi	1500
janng	800
libon	200

Chart Field : mlgreen
Chart Value : 7623.17
Number of Jobs : 27 (Total : 60 Percentage : 45.00%)
Average Number of Nodes/Procs : 1.96 / 9.93 (Total : 1641 / 3206 Percentage : 9.31%)
Average Runtime : 71.92 Hours (User Total : 1941.75 Percentage : 71.90%)
Average CPU Hour Production : 282.34 Hours (User Total : 7623.17)
Total CPU Hour Production : 31015.17 Hours (User Percentage : 24.58%)

The charts are built on a dynamic basis using PBS queue information that is updated every 15 minutes. There are two selections available for displaying the usage data, currently running jobs or currently queued jobs. The usage data can be grouped by user name, or group name. The subsequent statistics produced and displayed by hovering over the field are averaged based on the "group by" selection.

Description of chart field hover statistics:

Chart Field : Reports the current chart field name dependent on the "grouped by" selection (queue user name, or group name).
Chart Value : Reports the current chart value.

CCR Grid Computing Services: Grid Admin - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CCR University at Buffalo The State University of New York
Center for Computational Research GRID PORTAL
High Performance Grid Computing

PORTAL LOGOUT
User Tools
Manage Account
Grid General Info
Projects
Resources
Computational Grid
Job Submissions
Job/Queue Status
Data Grid
Data Grid Statistics
Network Status
Running/Queued Jobs
PBS Job History
Grid Portal Statistics
Center Fleet Statistics
User Information
Education/Outreach
Staff Only
CCR 8048

Chart based on: total CPU hours
usage data for: queued jobs
for: All resources
grouped by: user mlgreen

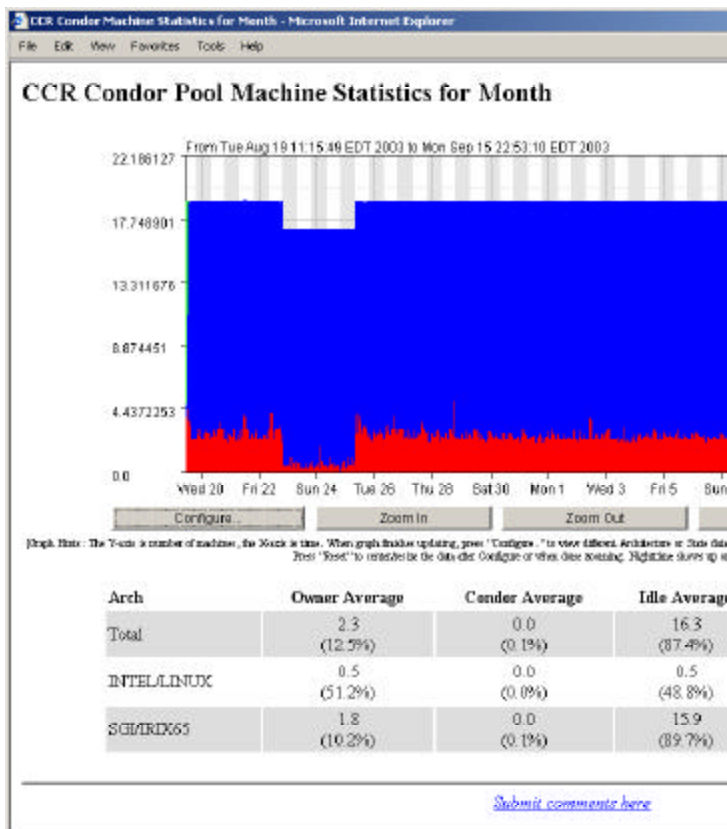
Job_Num	Job_Id	Username	Groupname	Queue	Nodes	Cnt	Proc	Cnt	Rank	Request_Time(hr)	Start_Time(hr)
1	23364	mlgreen	miller	grid	1	2	1	71.916666666667		queued	
2	23365	mlgreen	miller	grid	1	2	2	71.916666666667		queued	
3	23366	mlgreen	miller	grid	2	4	3	71.916666666667		queued	
4	23367	mlgreen	miller	grid	2	4	4	71.916666666667		queued	
5	23368	mlgreen	miller	grid	2	4	5	71.916666666667		queued	
6	23369	mlgreen	miller	grid	3	6	6	71.916666666667		queued	
7	23370	mlgreen	miller	grid	2	4	7	71.916666666667		queued	
8	23371	mlgreen	miller	grid	2	4	8	71.916666666667		queued	
9	23372	mlgreen	miller	grid	2	4	9	71.916666666667		queued	
10	23373	mlgreen	miller	grid	2	4	10	71.916666666667		queued	
11	23374	mlgreen	miller	grid	2	4	11	71.916666666667		queued	
12	23375	mlgreen	miller	grid	2	4	12	71.916666666667		queued	
13	23376	mlgreen	miller	grid	2	4	13	71.916666666667		queued	
14	23377	mlgreen	miller	grid	2	4	14	71.916666666667		queued	
15	23378	mlgreen	miller	grid	2	4	15	71.916666666667		queued	
16	23379	mlgreen	miller	grid	2	4	16	71.916666666667		queued	
17	23380	mlgreen	miller	grid	2	4	17	71.916666666667		queued	
18	23381	mlgreen	miller	grid	2	4	18	71.916666666667		queued	
19	23382	mlgreen	miller	grid	2	4	19	71.916666666667		queued	
20	23383	mlgreen	miller	grid	2	4	20	71.916666666667		queued	
21	23384	mlgreen	miller	grid	2	4	21	71.916666666667		queued	
22	23385	mlgreen	miller	grid	2	4	22	71.916666666667		queued	
23	23386	mlgreen	miller	grid	2	4	23	71.916666666667		queued	
24	23387	mlgreen	miller	grid	2	4	24	71.916666666667		queued	
25	23388	mlgreen	miller	grid	2	4	25	71.916666666667		queued	
26	23389	mlgreen	miller	grid	2	4	26	71.916666666667		queued	
27	23393	mlgreen	miller	grid	2	4	27	71.916666666667		queued	

Description of table fields:

Job_Num : Sequential counter for row number.
Job_Id : PBS job identification number corresponding to local PBS queue job number.
Username : PBS job owner local username.
Groupname : PBS job owner local primary group name.

ACDC-Grid Portal Condor Flock

CondorView integrated into ACDC-Grid Portal



CCR Center for Computational Research GRID PORTAL
High Performance Grid Computing

Condor High Throughput Computing

Condor Pool Statistics for CCR

Pool Resource (Machine) Statistics

- For the past hour
- For the past day
- For the past week
- For the past month
- For the month of [Jan] [Feb] [Mar] [Apr] [May] [Jun] [Jul] [Aug] [Sep] [Oct] [Nov] [Dec]

Pool User (Job) Statistics

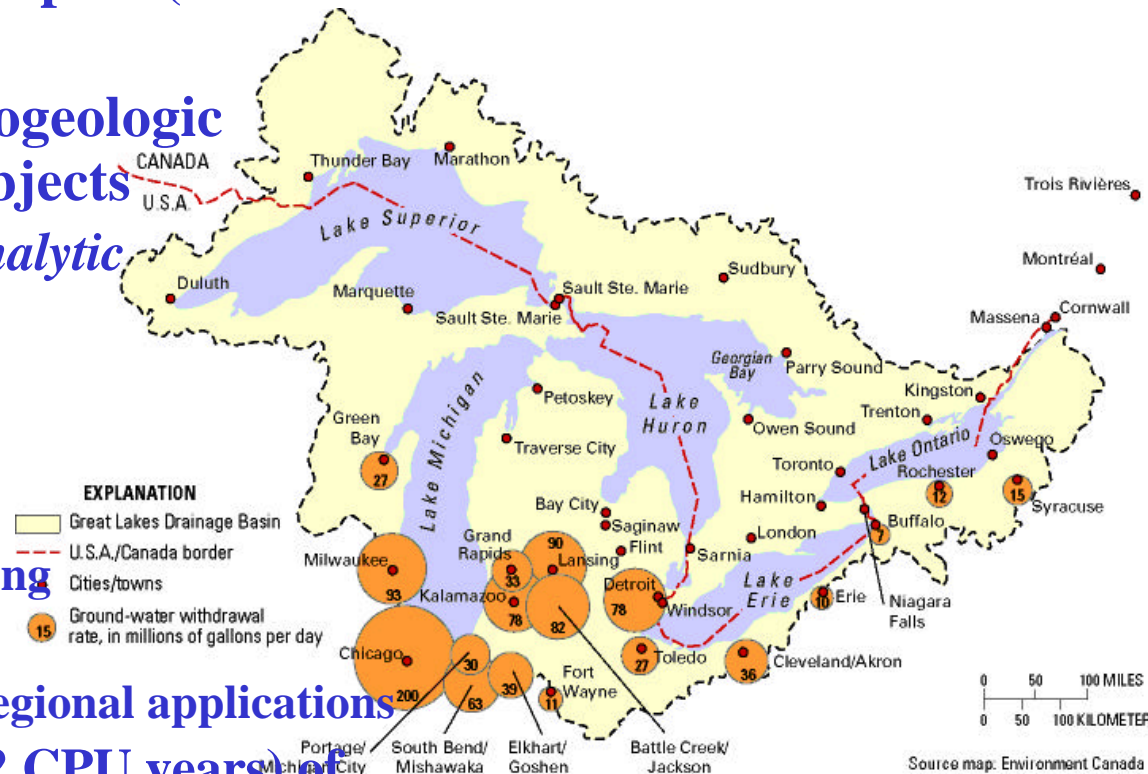
- For the past hour
- For the past day
- For the past week
- For the past month
- For the month of [Jan] [Feb] [Mar] [Apr] [May] [Jun] [Jul] [Aug] [Sep] [Oct] [Nov] [Dec]

[Submit comments here](#)

Advanced Center for Computational Research Data Center

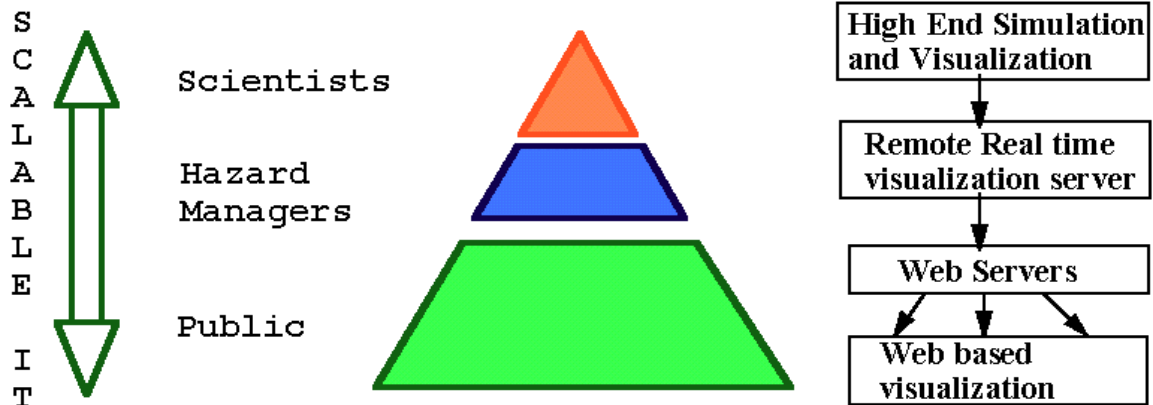
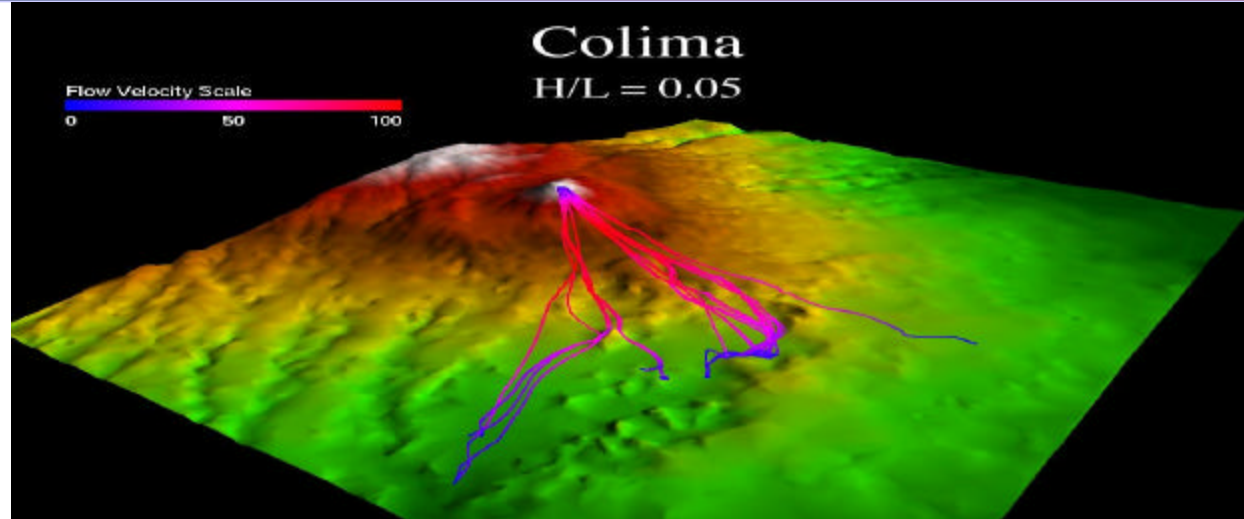
Groundwater Flow Modeling

- Regional-scale modeling of groundwater flow and contaminant transport (Great Lakes Region)
- Ability to include all hydrogeologic features as independent objects
- Current work is based on *Analytic Element Method*
- Key features:
 - High precision
 - Highly parallel
 - Object-oriented programming
 - Intelligent user interface
 - GIS facilitates large-scale regional applications
- Utilized 10,661 CPU days (32 CPU years) of computing in past year on CCR's commodity clusters



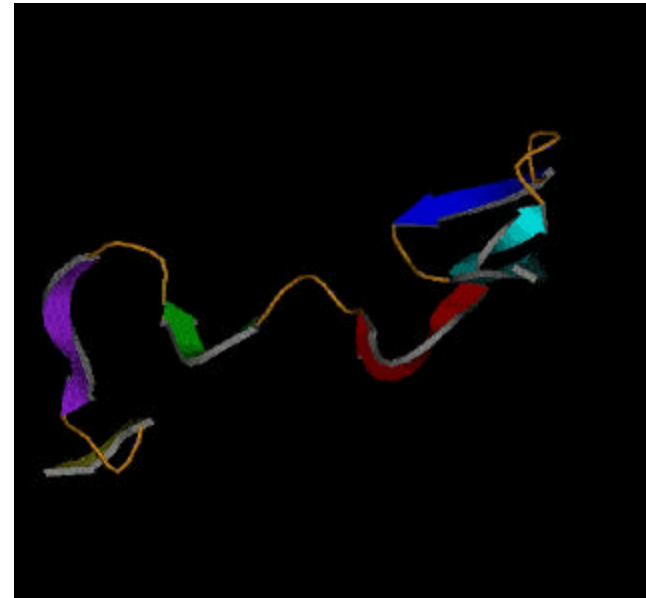
Risk Mitigation

- Integrate information from several sources
 - Simulation results
 - Remote sensing
 - GIS data
- Develop realistic 3D models of geophysical mass flows
- Present information at user appropriate resolutions



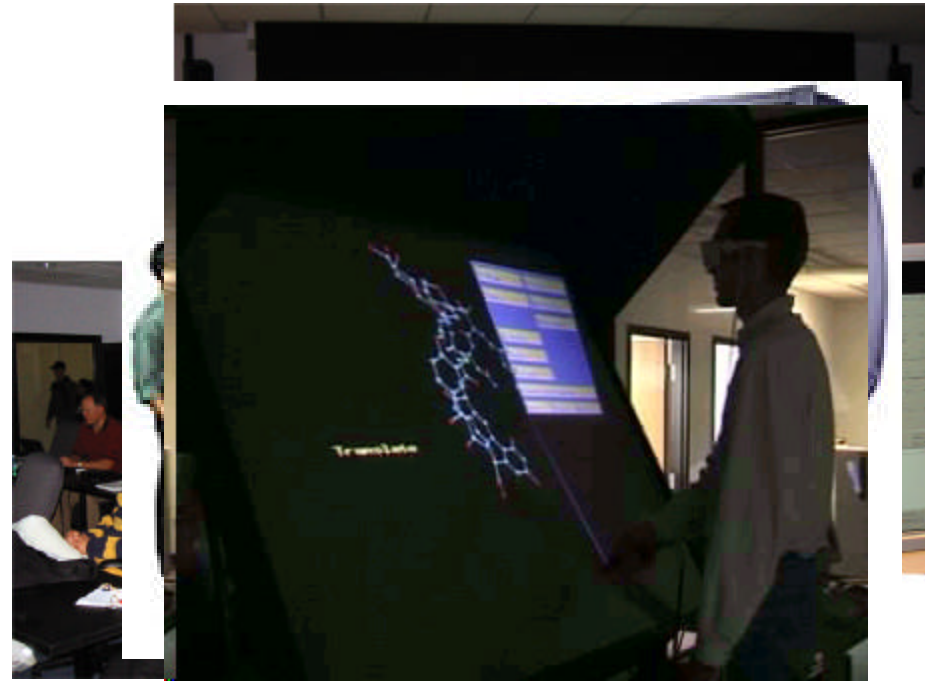
Protein Folding

- Ability of proteins to perform biological function is attributed to their 3-D structure.
- Protein folding problem refers to the challenge of predicting 3-D structure from amino-acid sequence.
- Solving the protein folding problem will impact drug design.



CCR Visualization Resources

- **Fakespace ImmersaDesk R2**
 - Portable 3D Device
- **Tiled-Display Wall**
 - 20 NEC projectors: 15.7M pixels
 - Screen is 11' ´ 7'
 - Dell PCs with Myrinet2000
- **Access Grid Node**
 - Group-to-Group Communication
 - Commodity components
- **SGI Reality Center 3300W**
 - Dual Barco's on 8' ´ 4' screen
- **VREX VR-4200 Stereo Imaging Projector**
 - Portable projector works with PC



CCR Visualization Projects

■ Peace Bridge Visualization Project

- ❑ CCR, Niagara College of Canada, Bergmann, IBC, eMedia Inc, Parsons Engineering

■ *StreetScenes*[®]

- ❑ Virtual Reality (VR) software solution for 3D visualization of surface traffic

■ Buffalo Niagara Medical Campus

- ❑ CCR, IBC Digital Inc.

■ Emergency response / GIS – Earthquake

- ❑ CCR

■ Williamsville Toll Barrier

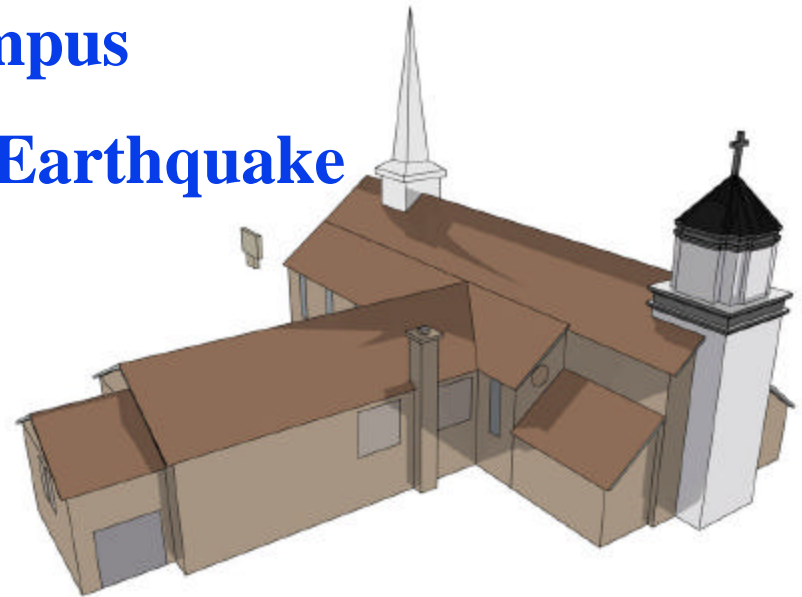
- ❑ CCR, TVGA

■ Accident Reconstruction

- ❑ CCR, TVGA

■ Biomedical Imaging

- ❑ CCR



Visualization in Planning Studies

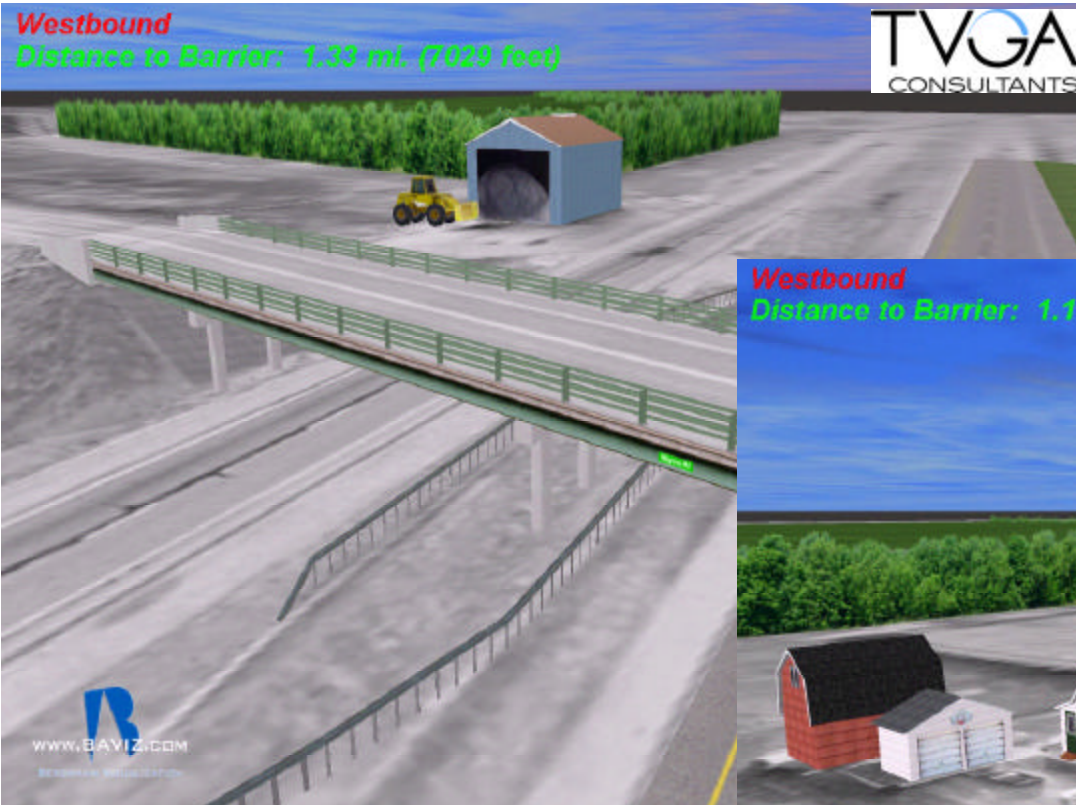


Williamsville Toll Barrier Improvement Project



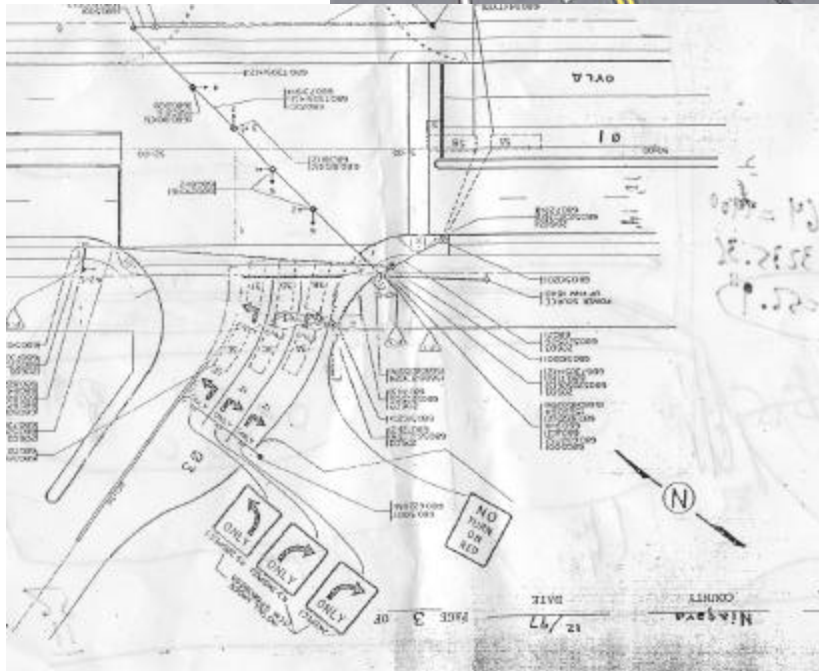
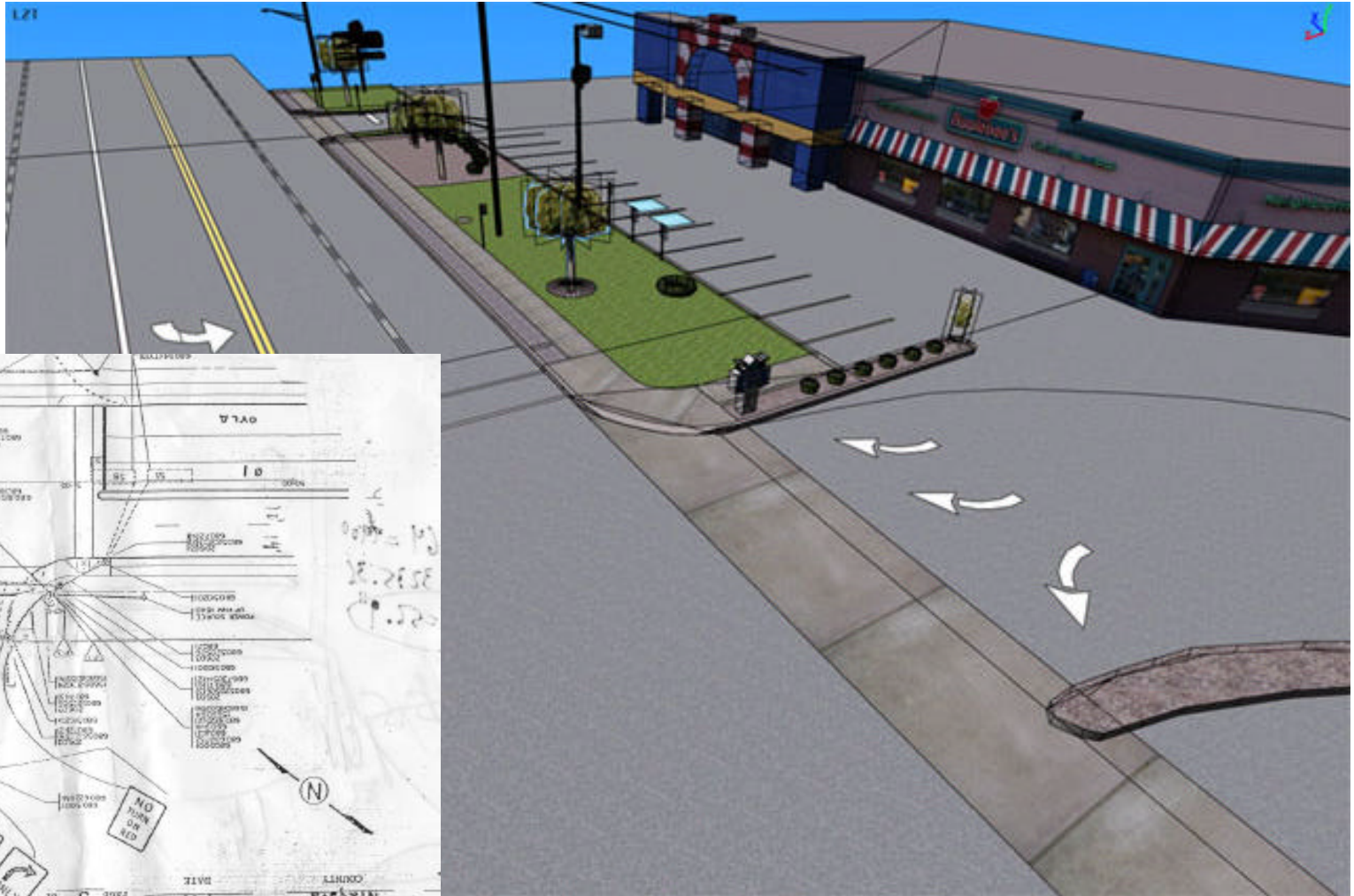
Initial Photo Match incorporating real and computer-generated components

Real-time Simulation



- Key Receptor Sites
- Multiple Viewpoints
- Fully Interactive
- Aerial Photography

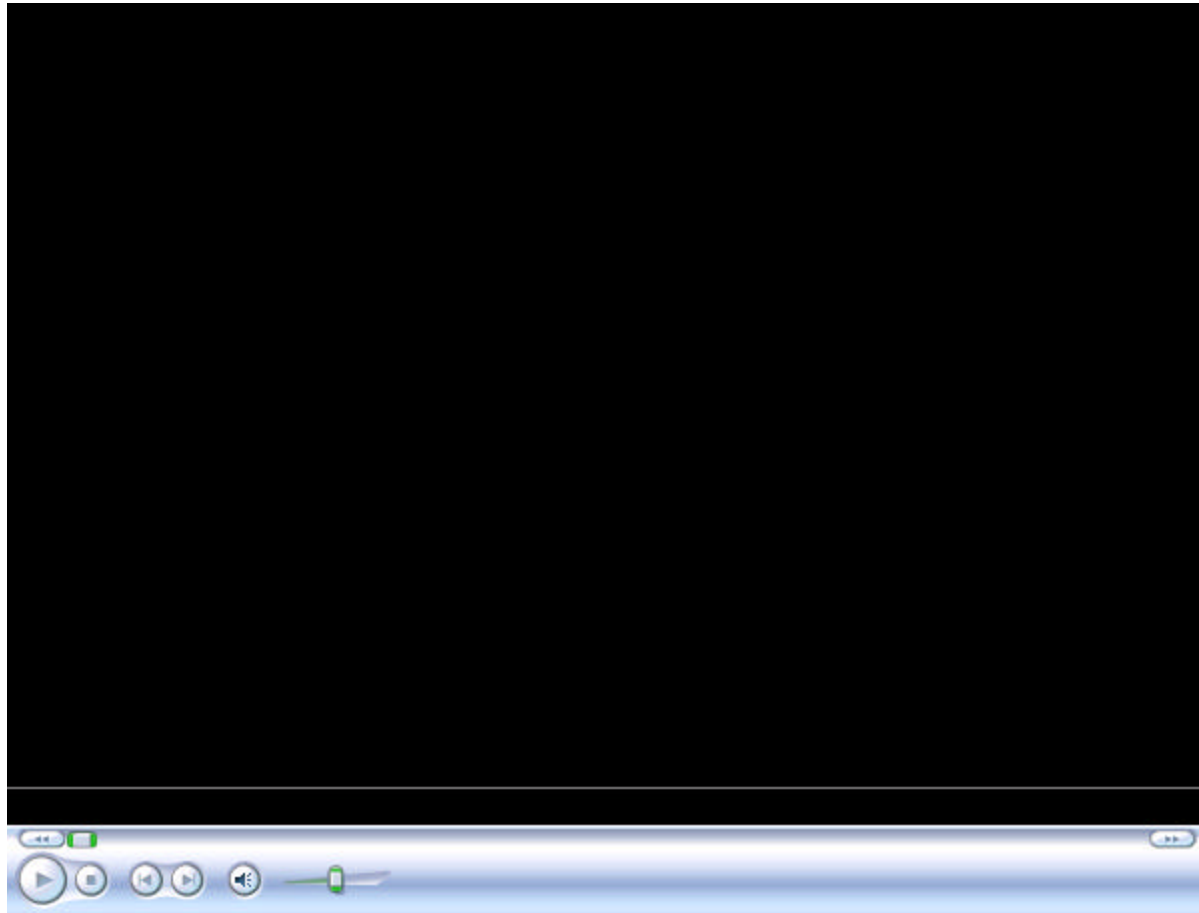
Accident Reconstruction



The Accident



Accident Animation (Driver's View)



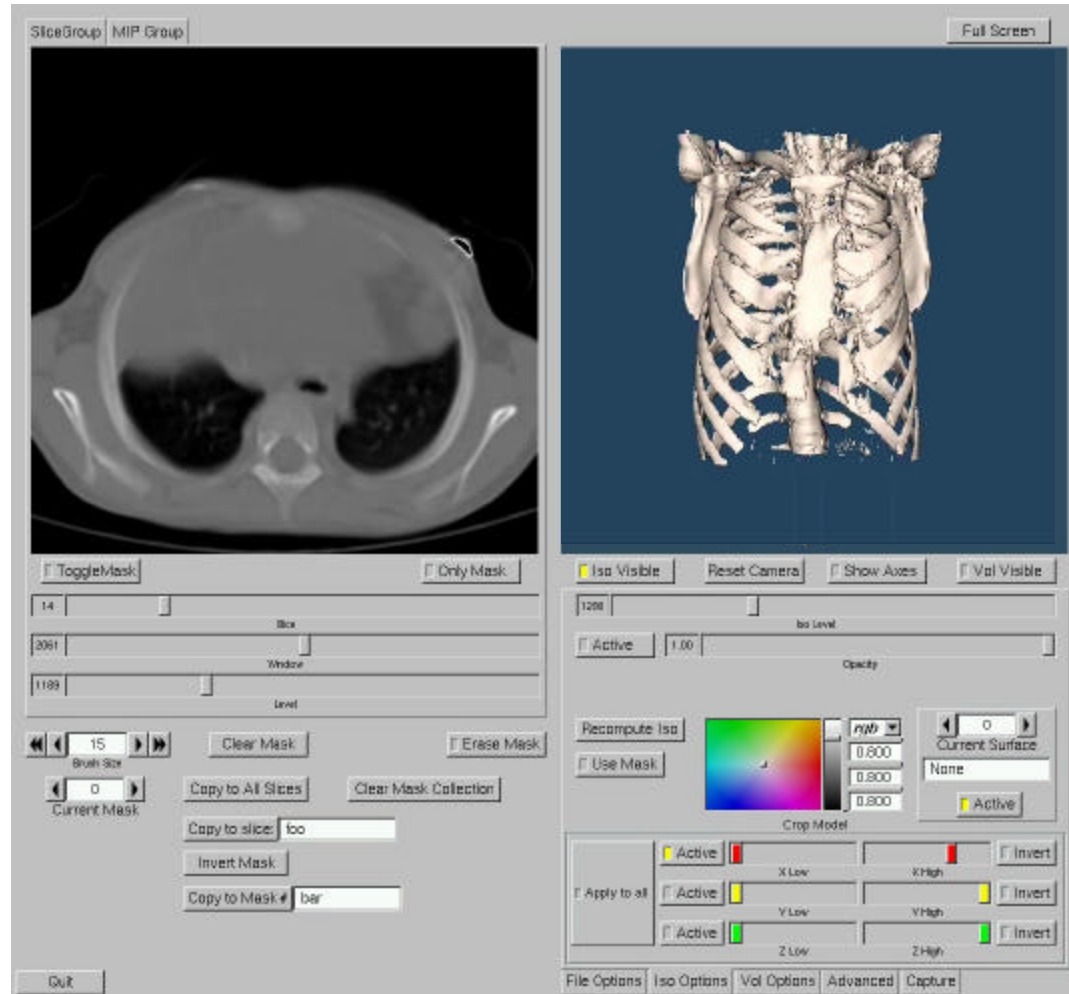
StreetScenes[®] Demo

- *StreetScenes*[®] is a Virtual Reality (VR) software solution for 3D visualization of surface traffic
- 3D model of proposed soccer stadium in Rochester
- Used *StreetScenes*[®] to import output file from Synchro traffic simulation



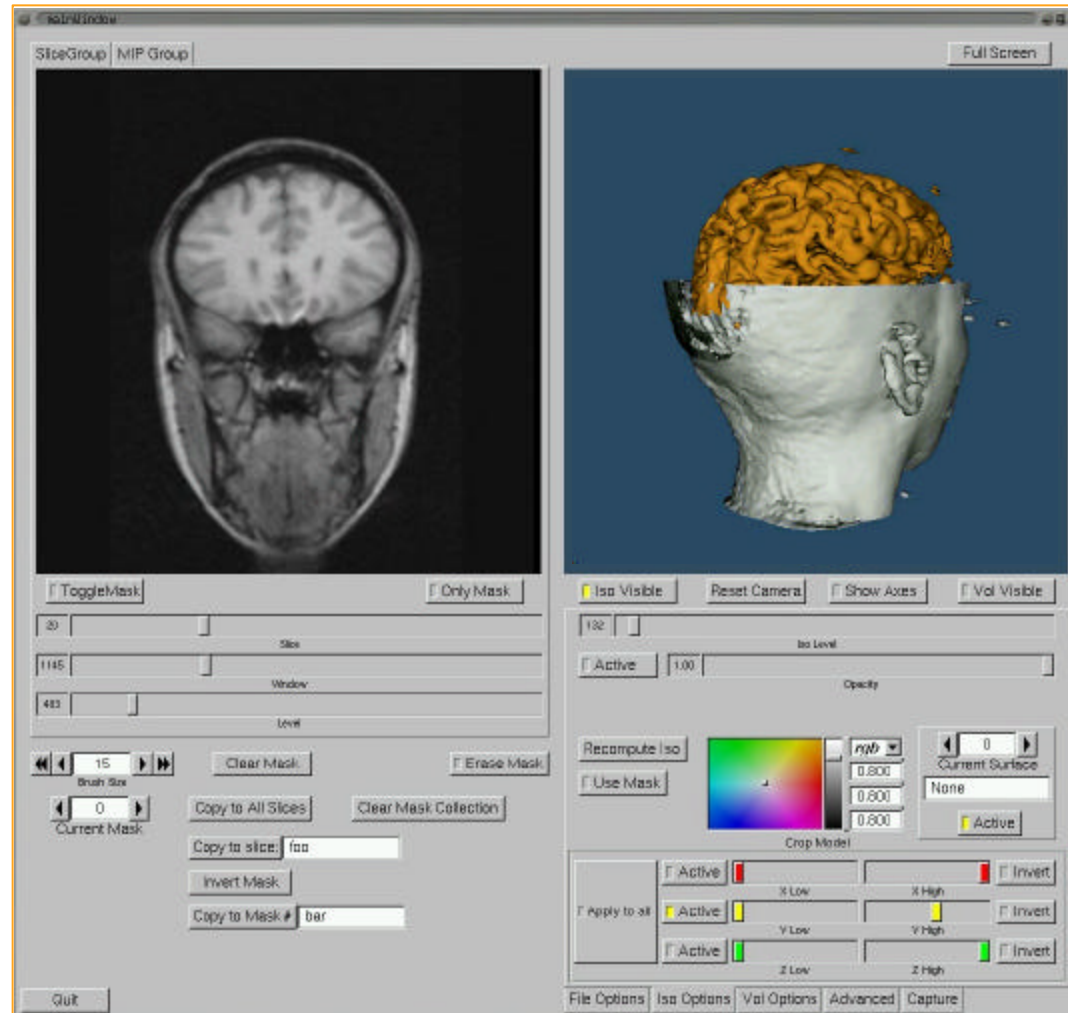
3D Medical Visualization App

- Collaboration with Children's Hospital
 - Leading miniature access surgery center
- Application reads data output from a CT Scan
- Visualize multiple surfaces and volumes
- Export images, movies or CAD representation of model



Multiple Sclerosis Project

- Collaboration with Buffalo Neuroimaging Analysis Center (BNAC)
 - Developers of Avonex, drug of choice for treatment of MS
- MS Project examines patients and compares scans to healthy volunteers



Select WNY Synergies

■ IBC Digital

- ❑ Gov. Pataki Visit
- ❑ Peace Bridge (Early & Current)
- ❑ Buffalo-Niagara Medical Campus
- ❑ Compute Cycles for Animation

■ Bergmann Associates

- ❑ Peace Bridge (Current)
- ❑ NYS Thruway Toll Plaza

■ Azar & More

- ❑ Reenactment of 1901 Pan Am Exhibition
- ❑ PHSCologram & Courses
- ❑ Avid Digital Editing

■ Niagara College

- ❑ Start up
- ❑ Peace Bridge (Current)

■ Hauptman-Woodward Medical Research Institute

- ❑ Computing
- ❑ Collaboratory

■ The Children's Hospital of Buffalo

- ❑ Medical Visualization

■ Veridian

- ❑ Battlespace Management

Outreach

- **HS Summer Workshops in Computational Science**
 - **Chemistry, Bioinformatics, Visualization**
 - **10-14 HS Students Participate Each Summer for 2 weeks**
 - **Project-Based Program**



Outreach

■ Pilot HS Program in Computational Science

- Year long extracurricular activity at Mount St. Mary's, City Honors, and Orchard Park HS
- Produce next generation scientists and engineers
- Students learn Perl, SQL, Bioinformatics
- \$50,000 startup funding from Verizon, PC's from HP



Media Coverage

20th Anniversary

BUFFALO Business First
Western New York's Business Newspaper

CALL US AT 202-253-3000

HMOs cut Medicare premiums

The new rules of Independent Health and United Healthcare will become effective March 1, pending presidential approval.

The change in its program to the Medicare standard appeal for Medicare Health Insurance.

In the past, Medicare HMOs have often been able to charge less than the standard Medicare rate. But the new rules will require HMOs to charge the same as the standard Medicare rate.

Patio home development proposed for Town of Aurora

The Town of Aurora is proposing a development in the town of Aurora, which would include a new residential development.

The project would include a new residential development in the town of Aurora, which would include a new residential development.

Grammy: Designs on Buffalo

Ani DiFranco, art director up for award

Ani DiFranco, art director up for award

Ani DiFranco, art director up for award

UB brings bioinformatics to a younger generation

Darcy Brown, a senior at Mount St. Mary Academy, learned about bioinformatics during a summer workshop at the UB Center for Computational Research.

Darcy Brown, a senior at Mount St. Mary Academy, learned about bioinformatics during a summer workshop at the UB Center for Computational Research.

Darcy Brown, a senior at Mount St. Mary Academy, learned about bioinformatics during a summer workshop at the UB Center for Computational Research.

BUFFALO NEWS

EDUCATION

An early look at bioinformatics

By EMMA D. SAPIRO
New Northtown Bureau

University at Buffalo undergraduate David Walsh works with Jaclyn Shaw, right, to demonstrate the "Next Generation Scientist" program. At left is Swarnan D'Avary.

computational science, it is being taught at Mount St. Mary, Orchard Park High School and City Honors School. About two dozen students are involved in the program, they work on smaller versions of the computers used at the research center.

Brown and the three other students in the program demonstrated and spoke about the program Thursday at Mount St. Mary. Awarding were officials from UB and Vermont, which funded the program with a \$50,000 grant.

"This is the first time in school, it's really not practical," Brown said. "Bioinformatics has shown me how to apply science in real life. It has really opened doors for me."

E. Bruce Pitzer, associate dean for research and sponsored programs at UB, said the program also is being implemented in high schools by developing a curriculum and training

students. It will expand into other schools in upcoming years.

The students met with a couple of selected teachers in their schools who also are receiving training and three UB undergraduate students.

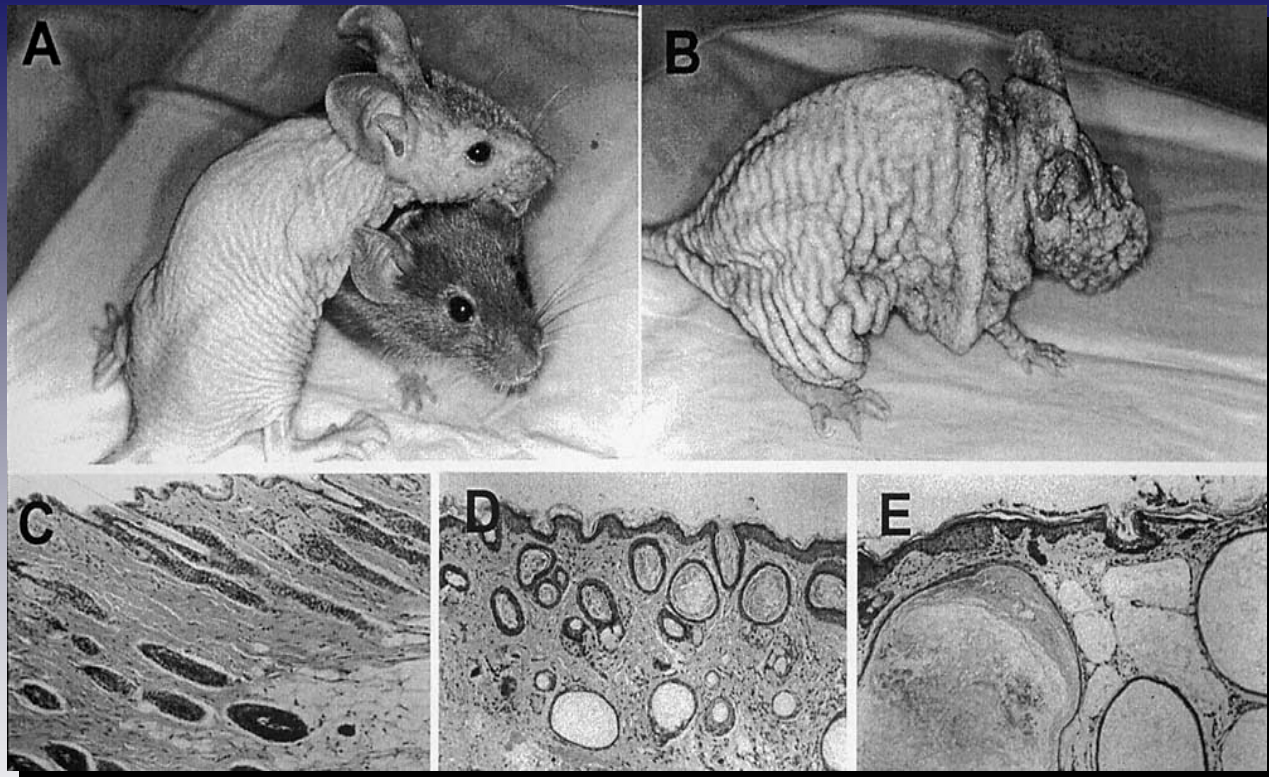
Senior Courtney Kowalski, who plans to major in mechanical engineering at Clarkson University, said bioinformatics has prepared her for her field of study. She said it's "going to give me a stronger background in engineering."

Because the students are all graduating, Brown said they are trying to recruit students for the program.

"Bioinformatics is really a different experience," she said. "You think of computers and computer programs and the way they are, and now you know the work that goes behind them."

emma.d.sapiro@ub.edu

Contact Information



miller@buffalo.edu
www.ccr.buffalo.edu