# Center for Computational Research

## All Hands Meeting - Tuesday, March 11, 2003


**Russ Miller, Director**

**Tom Furlani, Associate Director**


**"Top 10 Worldwide Supercomputing Center"**
**- www.gapcon.com**


**University at Buffalo**

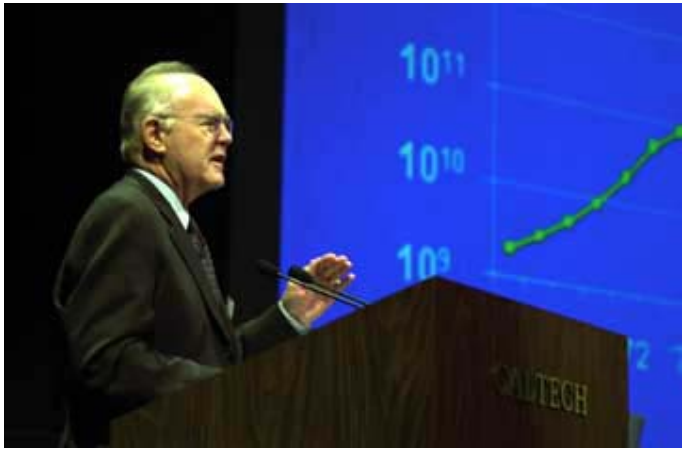
*The State University of New York*

# Agenda

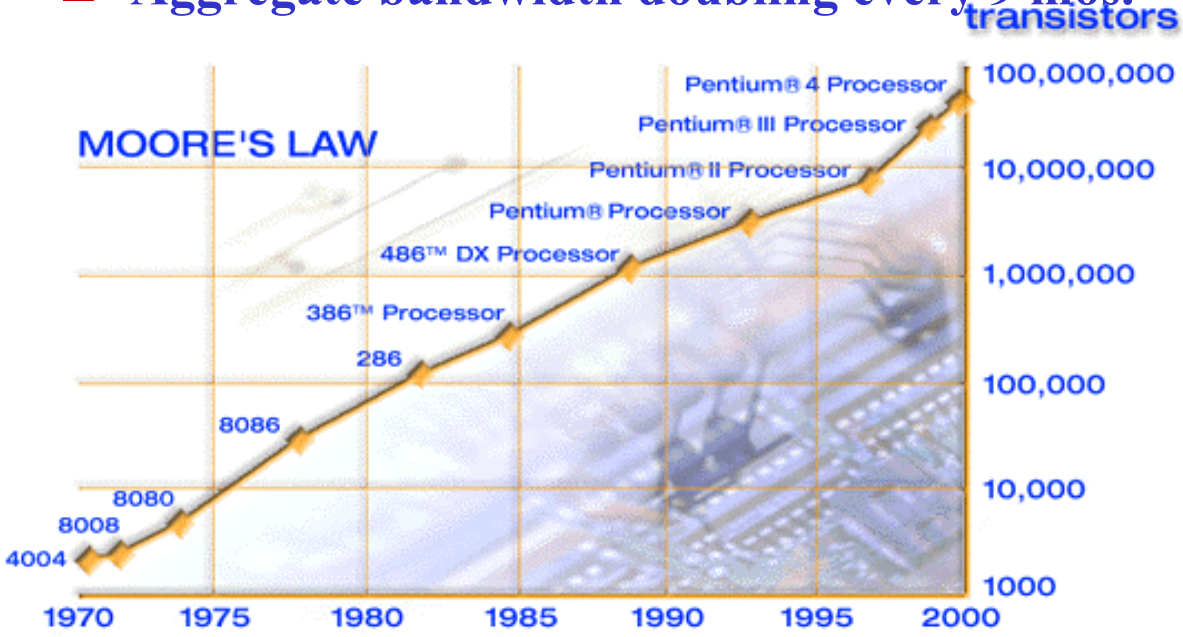
| | |
|---|---|
| 9:00 | Introduction - Miller |
| 9:15 | CCR Overview - Miller |
| 10:15 | Break |
| 10:30 | Overview of Projects at CCR - Furlani |
| 11:15 | Open Discussion/Networking |
| 12:00 | Lunch (on your own) |
| 1:00 | Intro to TDW and 3300W – Innus |
| 1:30 | Intro to AGN – Tilson |
| 2:00 | Parallel Computing at CCR – Jones/Tilson |
| 3:00 | Overview of Joplin - Cornelius/Green |
| 4:00 | Intro to Bioinformatics - Hu |

# 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

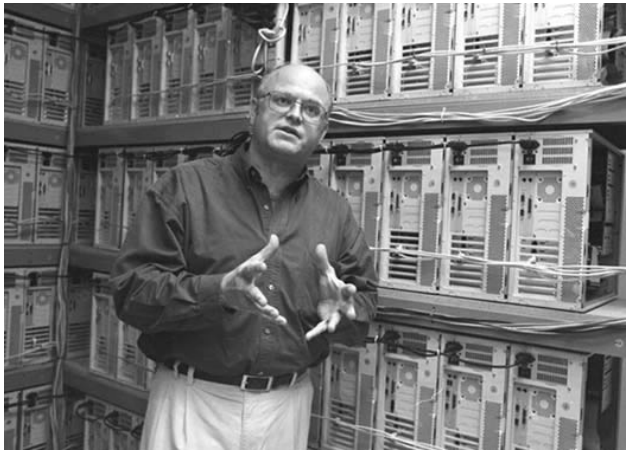


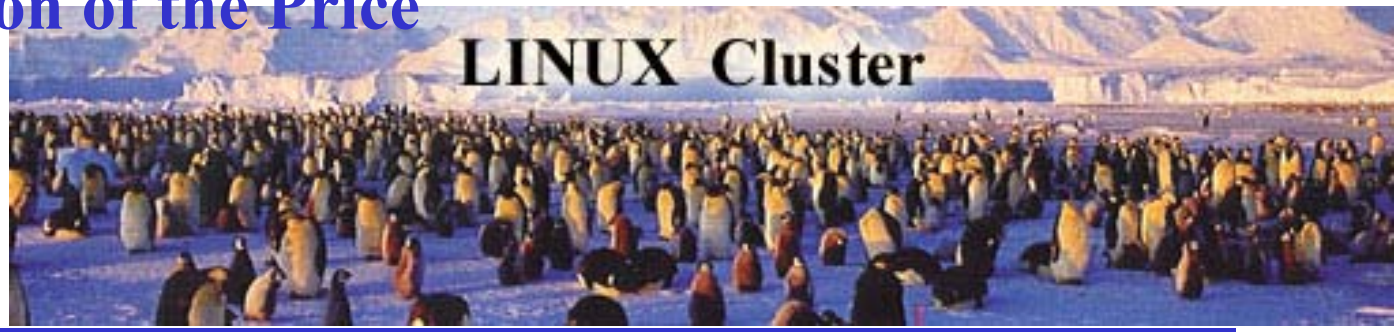

- A computation that took 1 year to run on a PC in 1985 would only take 5 mins to run on a PC today!
- A computation that runs in 2 hours on a PC today would have taken 24 years to run on a PC in 1985!

# What is a (Beowulf) Cluster?

- Industry Standard Hardware and Software
  - PC-Based Components (Intel or AMD)
  - Ethernet or Myrinet
  - Linux, PBS, MPI
  - "Commodity Off-The-Shelf" (COTS)
  - Introduced in 1995
- Operates as a Single System
- Rivals Performance of Traditional Supercomputer at a Fraction of the Price

Thomas Sterling
Caltech



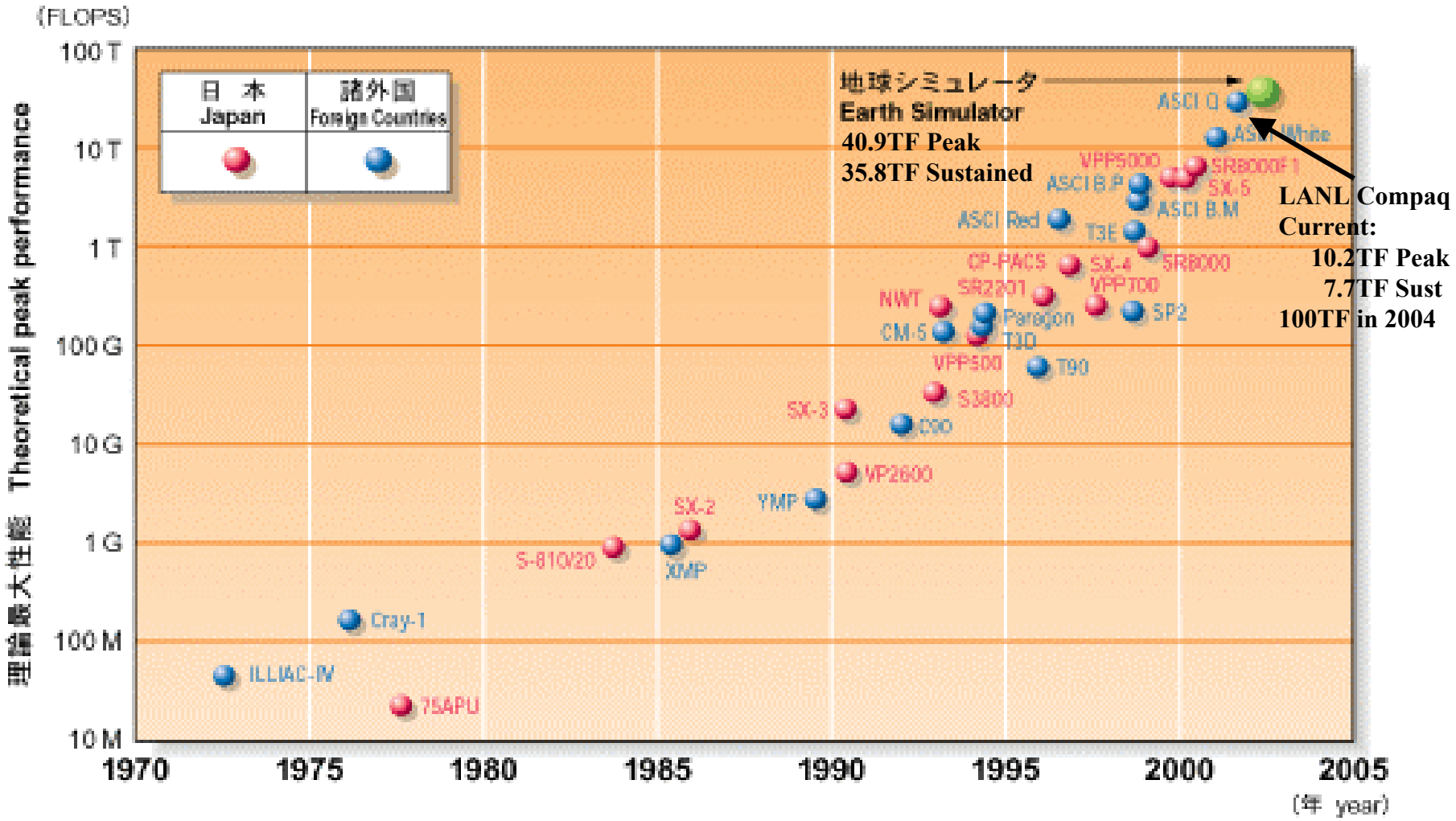


# Fastest Computers

| Year | Mach | Procs | GFlops |
|---|---|---|---|
| 1976 | Cray 1 | 1 | 0.1 |
| 1982 | Cray X-MP | 4 | 0.9 |
| 1986 | Cray 2 | 4 | 2 |
| 1989 | Cray Y-MP | 8 | 2.7 |
| 1989 | TMC CM-2 | 8192 | 28 |
| 1992 | TMC CM-5 | 1024 | 131 |
| A 1-year calc in 1980 = 5.4 sec today<br>A ~1990 HPC = a laptop today | | | |

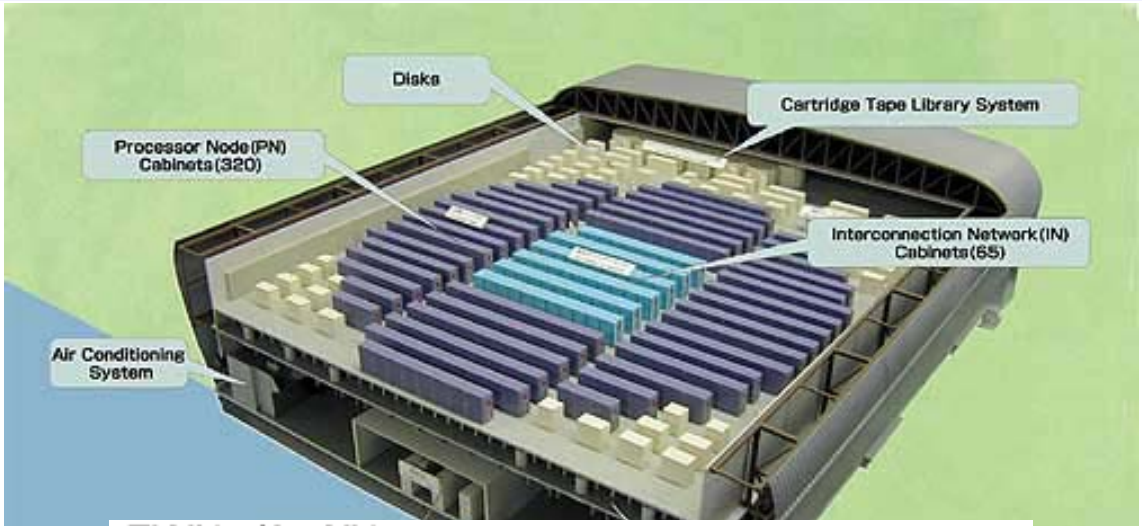
| Year | Mach | Procs | GFlops |
|---|---|---|---|
| 1993 | Cray T3D | 1024 | 152 |
| 1994 | Fujitsu VPP | 140 | 236 |
| 1996 | Hitachi SR2 | 2048 | 368 |
| 1997 | Intel ASCI-R | 9152 | 1830 |
| 1999 | SGI ASCI-BM | 6144 | 3072 |
| 2000 | IBM ASCI-W | 8192 | 12,288 |
| 2002 | NEC E.S. | 5120 | 40,960 |

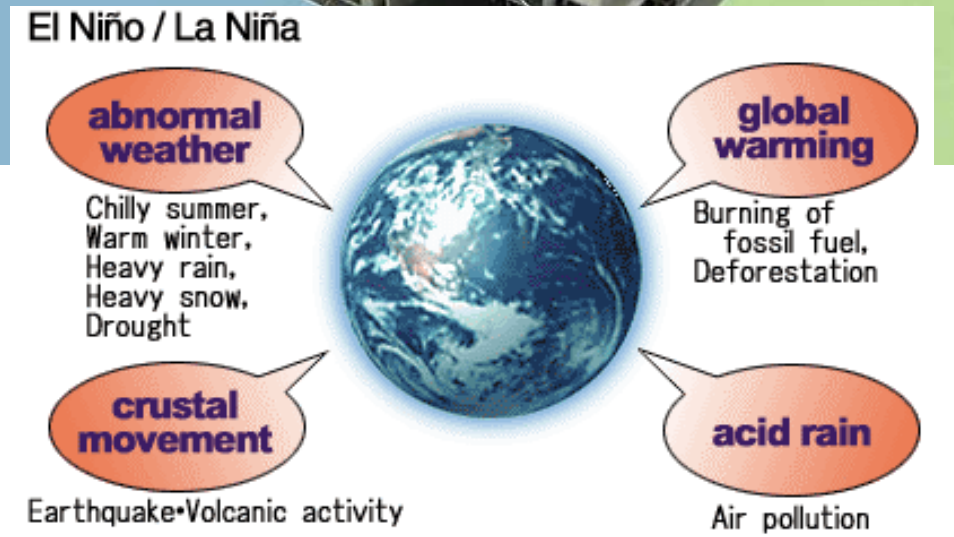
# Earth Simulator in Japan (NEC Vector Supercomputer)



# Earth Simulator

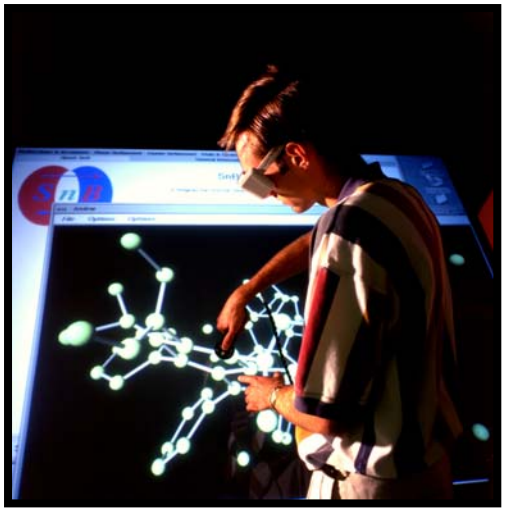
- 40TFlops Peak
- Homogeneous, Centralized, Proprietary, Vector
- Expensive!
- CFD-Weather, Climate, Earthquake
- 640 NEC SX/6 Nodes (5120 CPUs)
- Footprint = 4 tennis courts
- $6M/year in power



# Center for Computational Research

- **Facility Opened in May 1999**
- **Top Ten Worldwide Supercomputing Facility**
  - 9 Teraflops Aggregate Capacity (PR)
  - 3+TF Generally Available (Top 40)
- **Support Faculty-Based Research**
- **Foster High-Tech Research & Economic Development in WNY**
  - High-Performance Computing
  - High-End Visualization
- **Technology Transfer**
  - Seminars, Workshops and Tutorials



# General Information

- High-Performance Computing and High-End Visualization
  - 47 Active Research Groups in 27 Depts
- Leverage $300K NSF Grant ⇒
  - $41M Vendor Donations
  - $55M Grants & Contracts (non-bldg)
- Deliverables
  - 350 Publications & Pres., Hardware, Software, Algorithms
- Outreach (EOT)
  - Workshops, Training, Courses, Degree Programs
  - 25 Local Companies & Institutions
- Funding
  - NSF, NIH, NIMA, EPA, Keck, Sloan
  - NYS, SUNY, UB
  - IBM, SGI, Sun, DELL, HP (Compaq), Nortel, Myricom

# Staff

- Director – Dr. Russ Miller
- Associate Director - Dr. Tom Furlani
- Computational Scientists
  - Dr. Zihua Hu (Bioinformatics), Dr. Matt Jones (Physics), Dr. Jeff Tilson (Chemistry)
- Programmers
  - Steve Gallo (Grid Computing), Quoc Nguyen (Bioinformatics)
- Visualization
  - Adam Koniak (Urban Simulation), Martins Innus (Scientific Visualization)
- Systems Support
  - Cynthia Cornelius, Jon Bednasz, Sam Guercio, Tony Kew, Deborah Loke, Dori Macchioni , Jason Rappleye
- Office Staff
  - Brenda Sauka (Financial Manager), *Sally Elder (Secretary)*, Penelope Krebs (Receptionist)

# ExCom & AdvCom

- Executive Committee
  - Josephine Anstey, Media Study
  - Corky Brunskill, SENS
  - Philip Coppens, Chemistry
  - David Kofke, Chemical Engineering
  - Bruce Pitman, Math
- Internal Advisory Committee
  - Bruce Holm, Senior Vice Provost
  - Voldemar Innus, CIO
  - Mark Karwan, Dean School of Engineering
  - Bruce McCombe, Assoc. Dean, CAS & Director, CAPEM

# Space

- 1999
  - Norton 9: 2800 sq ft + 1100 sq ft Machine Room
- 2002
  - Norton 10 AGN, TDW: 600 sq ft
  - Norton 15 Machine Room: 1700 sq ft
  - Norton 118 Machine Room: 1700 sq ft
  - Norton 17 Office Area: 1600 sq ft
- Jan., 2003 - Delayed
  - Norton 5 & 7: 3000 sq ft
  - Provide Public Access to iDesk, AGN, Videoconferencing
  - Provide Appropriate Staff Office Space
  - Provide Visitor Office Space
- UPS System – Delayed from 1/02 to 3/03

# Making CCR Easier to Use

- Website Update
  - New Look/Feel by end of month
  - User Info/Hotpage - **Demonstration**
- ccr-help@ccr.buffalo.edu (RT-Tracker)
- Grid Software on the Horizon
  - Grid: flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources
  - Grid Portal Under Development
    - HPC Machines, Condor Flocks, Storage, Visualization, Instruments
    - Domain-Specific Areas of Portal
- Computational Chemistry Portal (ECCE)

# Machine Utilization

- Greater than 80% on most platforms
- Queuing Systems
  - Regularly adjusting queue structure to maximize throughput
  - 7 day queues discontinued
  - 3 day queue is now the longest
  - Throughput and utilization is significantly better
  - Feedback from users is welcome
  - Recent change under investigation: partition for parallel and sequential jobs on nash

# Outreach

- Collaboration with local institutions
  - Buffalo Neuroimaging Analysis Center – MS MRI Visualization
  - Children's Hospital – Medical Imaging
  - BuffLink – 3D Visualization of the Buffalo Niagara Medical Campus
  - Peace Bridge Authority
  - Bioinformatics in Buffalo (UB, Roswell Park, HWI)
  - H.S. Bioinformatics Clusters
  - Numerous Local Colleges
- H.S. Summer Workshop in Bioinformatics
- Tours
  - Faculty candidates, speakers, business leaders, politicians, funding agencies

# Computational Resources

- SGI Origin2000 → Origin3800 (Crosby): 1998+
  - 64 Processors (400 MHz)
  - Shared memory
  - 32 GB RAM
  - 400 GB Fibre Channel Disk
- Full Maintenance for 2 years
- Primary Applications
  - Computational Fluid Dynamics
  - Quantum Chemistry (ADF/Jaguar/Q-Chem)
  - Computational Mechanics
  - Physics

# Computational Resources

- IBM RS/6000 SP (Stills): 1998+
  - 78 Processors
  - 26 GB RAM
  - 640 GB Disk Storage
- Primary Applications
  - Computational Fluid Dynamics (CFX4)
  - Chemical Engineering (Molecular Dynamics, etc)
  - Quantum Chemistry (NWChem, ADF, Columbus, CHARMM, Q-Chem)
  - MAE (PETSC)
- Off Maintenance Jan 2004
  - *Requires Exit Strategy*

# Computational Resources

- Sun Microsystems Cluster (Young): 1999+
  - 48 Sun Ultra 5s (333MHz)
  - 16 Dual Sunblades (750MHz)
  - 30 GB RAM, Myrinet
  - Sun Solaris OS
- Large Number of Sequential Jobs
- Primary Applications:
  - Quantum Chemistry (ADF)
  - Chemical Engineering
  - CHARMM
- Processors Showing Age
  - Redeploy when Opportunity Presents

# Computational Resources

- SGI Intel Linux Cluster (Nash): 2000
  - 150 PIII Processors (1 GHz)
  - 75 GB RAM, 2.5 TB Disk Storage
  - Myrinet2000
- Primary Application Areas:
  - Chemical Engineering
  - Computational Fluid Dynamics
- Off Maintenance in 1.5 years
- Stability Issues due to Poor Hardware
- Consider Exit Strategy

# Computational Resources

- New Dell Linux Cluster (Joplin): 2002
  - 600 P4 Processors (2.4GHz)
  - Myrinet2000
  - 600 GB RAM; 40 TB Disk
- 22nd Fastest Computer in the World
- Large Production Machine
- Highly Parallel Codes
- Primary Application Areas:
  - Chemical Engineering
  - Computational Mechanics
  - Computational Fluid Dynamics
  - Environmental Chemistry
  - Quantum Chemistry (ADF,NWChem)


# Computational Resources

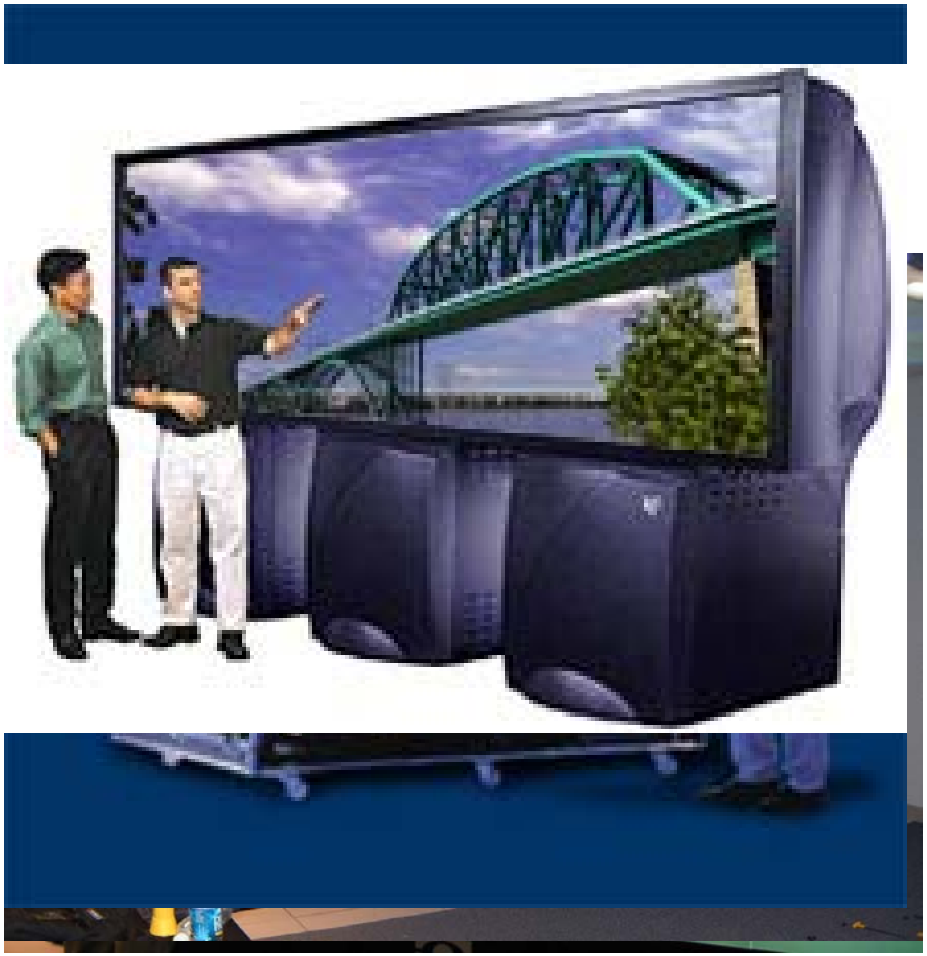
- Sun Enterprise 6000 (Database): 1998
  - 4 UltraSparc-II 336 MHz
  - 2 GB RAM
  - 38 GB disk + 2 × T3 Disk Arrays
- Primary application areas:
  - Large Oracle Database Application
  - Data Mining – Tops Friendly Markets

# Computational Resources

- Apex Bioinformatics System: 2003
  - Sun V880 (3), 6800, 280R (2), PIIIs; Sun 3960: 7 TB Disk Storage
- BioACE (Bioinformatics Application Computing Environment)
  - Available Directly from CCR Website
  - Faculty/Staff/Students from UB, RPCI, HWI
  - Home to Many State-of-the-Art Bioinformatics Software Packages
    - DNA and Protein Sequence Analysis (on hold)
    - Database Search
    - Gene Expression Profiling
    - Data Visualization
- Missing Small Linux Cluster (DNA & Prot. Seq. Anal.)
- GCG Implementation Postponed
- Workshop Tomorrow on InforMax Packages

# Visualization Displays

- FakeSpace Sys ImmersaDesk
  - Portable
- 3300W Visualization Display
  - 2 Channel Superwide Display – Dual Barco 808 Projectors
- Tiled Display Wall
  - 2D
  - 11.5' x 6.5' image
  - Myrinet Interconnect
  - 20 projectors/Dell 530's
  - Resolution: 5105 x 3072 pixels (15M Pixels)



# Visualization Computing

- SGI Onyx 300 – 10 Processors
  - Infinite Reality 4 Graphics Pipes (1GB Texture memory)
  - Used to drive the iDesk and the 3300W simultaneously
- SGI Onyx2 – 6 Processors
  - Infinite Reality 2 Graphics Pipes (64MB Texture memory)

# Access Grid Node (AGN)

- Facilitates group to group collaboration

# Budget (1)

- Personnel: $1.2M/yr by UB
- Discretionary Operating: ≤ $100K/yr by UB
  - Travel, training, equipment, software, maintenance, memberships, repairs, tapes, supplies, phones, installs, shipping, promotional material, networking, furniture, etc
- Equipment Acquisitions: Ad Hoc
  - Approx. $6.5M to date (4 year period)
- Comparison

| | Budget/year | Staff |
|---|---|---|
| CCR | $1.3M | 18 |
| NCSC | $7.6M | 23 |
| OSC | $6.6M | 48 |
| U Minn | $7.0M | 24 |

# Budget (2)

- CCR-Leveraged Grants & Contracts
  - CCR Grants & Contracts: $12M
  - Vendor Donations (in-kind): $41M
  - External Research Funding: $43M
    ($290M incl corporate, construction funds, foundation support, etc.)
- User Support
  - Critical that CCR is included in routing sheet for proposals that will use facilities
  - Direct costs for CCR are very important in these financial times
- Student Support from Operating, Workstudy, or External Grants (including REUs)

# Charge-Back Scheme

- Schemes for Cost-Recovery Under Consideration
- Basic Rates Have Been Determined
  - SGI Linux: $0.4 per processor per hour
  - SGI Origin: $0.6 per processor per hour
  - IBM SP: $1.0 per processor per hour
  - iDesk $69.0 per hour
- What Does This Mean for Users?
  - $1,000 - $22,000 per month
- Status of Charge-Back Scheme Unknown



# Visualization Review

- External Panel Convened to Review Visualization Efforts on Campus
- Site Visit on Dec 19th, 2002
- Committee did not read Documents
- No Debriefing with Committee
- Brief Conversation with Provost in Feb.
- Positive Feedback on CCR
- Central Funding Should be Provided
- Charge-Back Model is Not an Option
- Formation of Provost Advisory Panel for CCR

# Consolidation & Upgrade Plan

- Retain
  - Dell Linux P4 Cluster
- Acquire
  - 1TF SMP System (Regatta, Alpha, Itanium)
  - 0.5TF General Purpose PC farm
  - Storage Area Network (SAN)
- Evaluate
  - Sun6000 Database System
- Retire/Redeploy
  - IBM SP
  - Sun Cluster
  - SGI Linux Cluster
- Derailed with Cancellation of $6.5M HP EV7 & Intel Systems (PC Farm & BioCCR System)

# Future Plans

- Finances Restrict Planning
- Needs Assessment
  - Stable and Reasonable Operating Budget
    - Level Similar to NCSC, U Minn, OSC
    - Will Allow for Strategic Planning
    - Constant Refresh of Equipment
    - Travel, Training, Workstations, Supplies
    - Software, Tapes, Cables
- *???Administrative Discussion???*

# Break

- Relax, Converse, Food/Coffee

| | |
|---|---|
| 10:30 | Overview of Projects at CCR - Furlani |
| 11:15 | Open Discussion/Networking |
| 12:00 | Lunch (on your own) |
| 1:00 | Intro to TDW and 3300W – Innus |
| 1:30 | Intro to AGN – Tilson |
| 2:00 | Parallel Computing at CCR – Jones/Tilson |
| 3:00 | Overview of Joplin - Cornelius/Green |
| 4:00 | Intro to Bioinformatics - Hu |