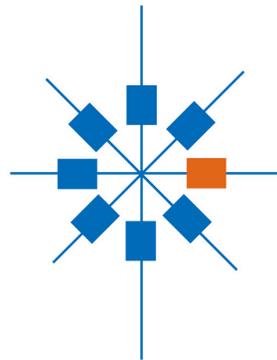


INFORMATION TECHNOLOGIES
& SERVICES DIVISION

ITSD Infrastructure and Services

**Tammy Welcome, Deputy for Infrastructure,
Scientific Cluster Support**

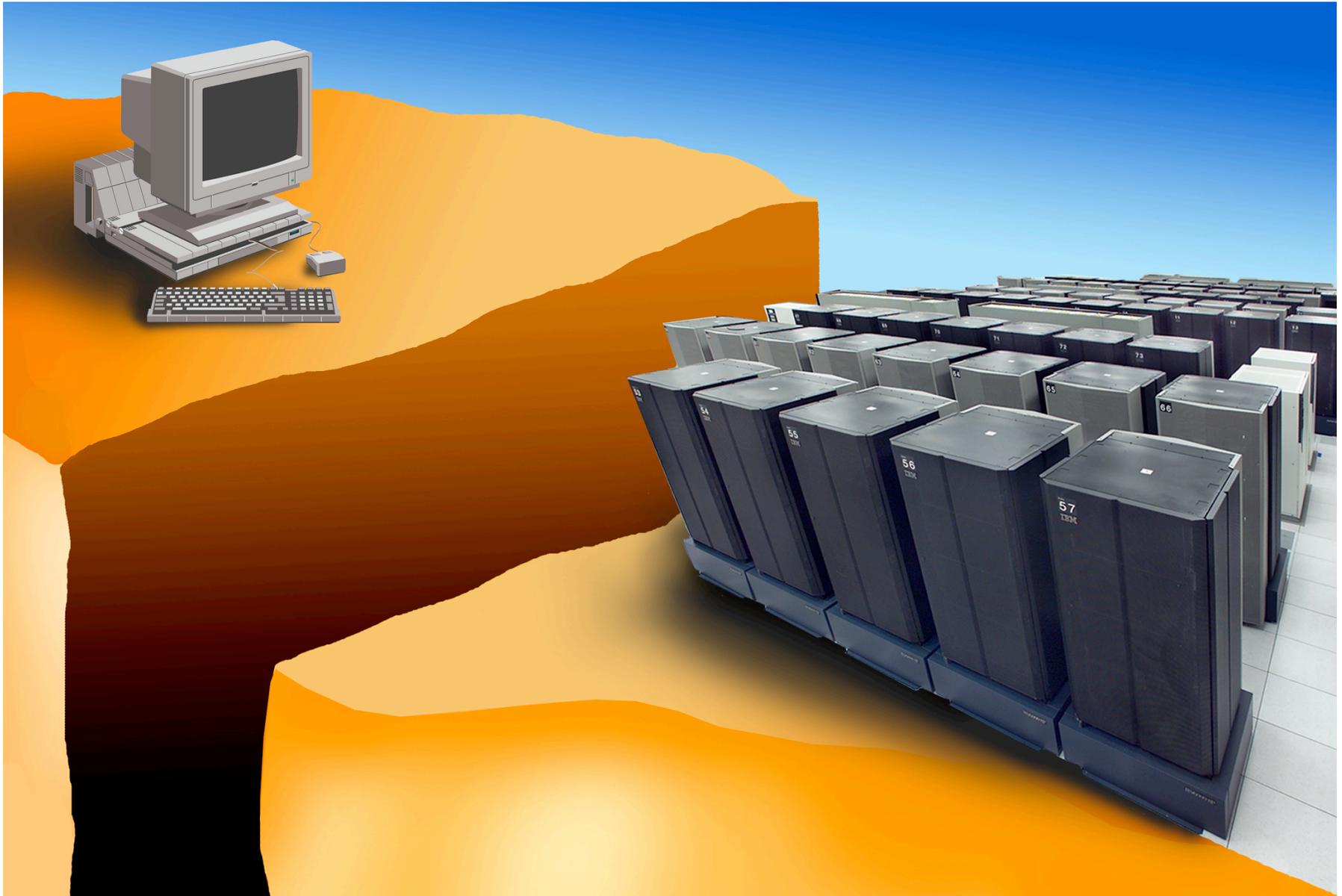




SCS
Scientific Cluster Support

- **An emerging and unmet need**
- **Finding a solution**
- **Status of our efforts**
- **Challenges still ahead**

Midrange Computing Gap



- **Charter**
 - Assess current computing needs
 - Identify common areas and solutions
- **Findings**
 - Growth in number of clusters on site due to affordable hardware
 - Setting up and running efficient cluster is not trivial and not cheap
 - Some clusters not in machine rooms
 - System administration by students, postdocs, and PI
 - Resource and time outlays in recovering from security hacks
 - Outages due to incorrect system configuration
- **Recommendation: ITSD provide support services for individual clusters**

- **Four-year program started January 2003**
- **Science projects purchase their own Linux clusters**
- **ITSD provides full cluster support**
- **Stage clusters over two years**
 - **six clusters in FY03**
 - **four more clusters in FY04**

ITSD provides Linux cluster support services

- Pre-purchase consulting**
- Procurement assistance**
- Computer room space**
- Initial setup and configuration**
- Systems administration**
- Cyber security**

- **FY03**
 - Enhance technical expertise
 - Identify software and support infrastructure
 - Form steering committee
 - Install six clusters
- **FY04**
 - Move four additional clusters into project
 - Provide ongoing support for 10 clusters
- **FY05**
 - Develop path forward recommendation
- **FY06**
 - Implement path forward
 - Wrap up four-year project

- **Provide effective tools for scientists**
- **Use proven approaches that enable us to provide production capability**
- **Adopt approach scalable from one cluster to several clusters**

- **Many options in building and configuring a cluster**
- **Allow freedom of choice for system features that matter most to scientists**
 - **specific hardware vital statistics (e.g. MHz, memory size, etc)**
 - **end user specific configuration and software**
- **Standardize**
 - **basic hardware setup**
 - **software for system management and administration**

Chemical Sciences	PI: Miller Contact: Michele	Semiclassical Molecular Reaction Dynamics: Methodological Development and Application to Complex Systems	New 30 Node Linux Cluster
Chemical Sciences	PI/Contact: M. Head-Gordon	Parallel electronic structure theory	New 15-25 Node Linux Cluster
Chemical Sciences	PI: Lester Contact: Aspuru- Guzik	Quantum Monte Carlo for electronic structure	Existing 19 Node Linux Cluster
Materials Sciences	PI: Arup Chakraborty Contact: Hagan, Dinner	Signaling and Mechanical Responses Due to Biomolecular Binding	New 50 Node Linux Cluster
Physical Bioscience	PI/Contact: Kim/Adams/ Brenner/Holbrook	Structural Genomics of a Minimal Genome Computational Structural & Functional Genomics A Structural Classification of RNA Nudix DNA Repair Enzymes from <i>Deinococcus radiodurans</i>	Existing 32 Node IBM Linux Cluster to be expanded to 64 Nodes
Environmental Energy Technologies	PI: Gadgil/Brown Contact: Finlayson/Tonse	Airflow and Pollutant Transport in Buildings Regional Air Quality Modeling Combustion Modeling	New 20 Node Linux Cluster
Earth Sciences	PI: Hoversten/Majer Contact: Daley	Geophysical Imaging	Existing 16 Node Linux Cluster to be expanded to 24 nodes
Life Sciences	PI/Contact: Eisen	Computational Analysis of cis-Regulatory Content of Animal Genomes	Existing 6 node Linux Cluster to be expanded to ~ 20 nodes
Life Sciences	PI: Cooper/Tainer Contact: Frankel	Protein Crystallography and SAXS data Analysis for Sibyls/SBDR	New 16 Node Linux Cluster
Physics	PI/Contact: White	Computing the Universe	Existing 10 Node Linux Cluster to be expanded

PI: Chakraborty
Materials Sciences

84 AMD 2200+ MP processors
42 GB aggregate memory
1 TB disk storage
Fast Ethernet interconnect
302 Gflop/s (theoretical peak)



PI: Gadgil/Brown/Tonse
Environmental Energy Technologies

24 AMD 2400+ MP processors
24 GB aggregate memory
2 TB disk storage
Myrinet interconnect
96 Gflop/s (theoretical peak)



PI: Hoversten/Majer
Earth Sciences

14 1 GHz Intel Pentium III processors
16 1.4 GHz Intel Pentium III processors
30 GB aggregate memory
800 GB disk storage
Myrinet interconnect
72 Gflop/s (theoretical peak)

PI: Miller
Chemistry

<currently in planning phase>

- **Feedback from first cluster group – very satisfied**
- **Warewulf cluster system configuration and management tools developed by ITSD gaining acceptance in community/campus**
- **Additional interest in cluster support services**
 - **Berkeley Center for Structural Biology**
 - **NSD Greta (Gamma Ray Energy Tracking Array) Detector Project**
 - **Molecular Foundry**

- **Securing systems**
- **Providing support for user application development**
- **Managing large systems**
- **Charting path forward**
 - **Partner with Computing and Communications Services Advisory Committee (CSAC) and project participants**