

# Scientific Computing Support Program

---



## Proposal from ITSD

### Result of Workshop and Subsequent Discussion with Individual Scientific Projects

Presented to DDM

October 30, 2002

Sandy Merola

# Scientific Computing Support Program Outline



- **Program Defined**
  - Program provides services needed and requested by scientific Divisions
  - Leaves computer ownership with the Scientific Projects
- **Goals/Benefits**
  - Improve Scientific Computing utilization by ensuring professionally configured and administered systems
  - Allow scientists to do science (and not systems administration)
  - Minimizes overall costs from a Laboratory perspective
- **Potential Program Scope and Cost**
  - Three year program, including 7 divisions, for up to 10 projects
  - Three approaches: Full Support, Cost Sharing, Minimal “Start Up”
  - Cost of support program ranges from \$605K - \$780K over three years
  - Minimal “Start up” only program is \$380K over 3 years

# Program Defined



## **Provides services needed and requested by Divisions**

- **Computer room space (9 of 10)**
  - Proper space, cooling, fire protection essential for high density clusters
  - Major barrier for many projects
- **Test cluster (3 of 10)**
- **Pre-purchase consulting (7 of 10)**
- **Initial setup and configuration (8 of 10)**
- **Cybersecurity (9 of 10)**
- **System admin (9 of 10)**
- **Coordination of Center of Excellence (7 of 10)**
- **Backups available at additional cost**

## **Leaves computer ownership with Scientific Projects**

# Goals and Benefits



## Improve computing utilization by ensuring professionally configured and administered systems

- Purchased, set up and configured properly
- Less downtime, better reliability, more stable
- Improved utilization
- Secure

*“The expert advice and assistance available through LBNL in choosing and deploying new security features would allow us to significantly **improve our productivity and increase our use of the cluster...** We have **spent considerable resources last year recovering data and deploying additional security features as a result of a single hacker**” – Berkeley Structure Genomic Center*

*Current cluster is expensive to maintain ourselves and **would free up the time spend “learning” when we have system problems.** Have had **several long outages in the past 6 months due to incorrect system config.** One problem resulted in a loss of **0.5 TB** of data – Martin White, Physics Dept.*

# Goals and Benefits



**Allow scientists to do science (and not systems administration)**

**Ability to redirect resources to science endeavors**

*“I (the PI) do it all. I am the only one in the lab with any sysadmin skills and I perform all of the sysadmin tasks” – Michael B. Eisen, Life Sciences*

**For some, incentive to make a desired change to cluster computing**

*It is essential to begin developing performance algorithms geared towards clusters in order to be state of the art in electronic structure in the future - Martin Head-Gordon, Chemical Sciences Division*

# Minimizes Overall Costs from a Laboratory perspective

---



## **Economy-of-scale cost savings**

- Leveraging of technical expertise
- Standardized approach to support saves \$\$

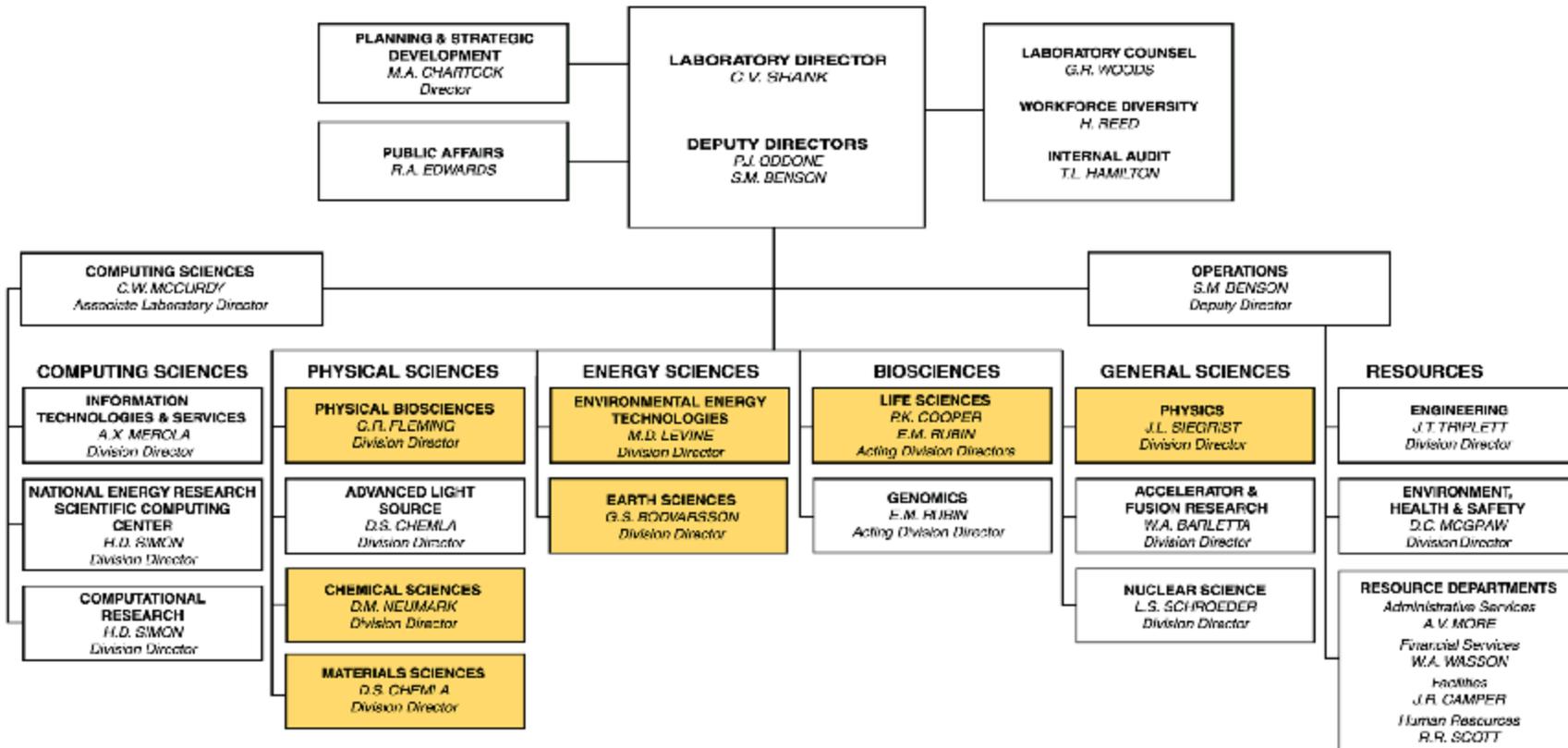
## **Recovers Lost Opportunities**

- Allows scientists to do science instead of systems admin
- Minimizes guesswork in purchase, configuration decisions
- Results in more computational throughput

## **Encourages scientific advantage of computing**

- Motivates paradigm shift to cluster and parallel computing
- Access to scientific computing helps keeps our science competitive with science at other labs

# Program Would Benefit LBL Science Divisions



07/25/02

# List of Proposed Applicants



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
Material 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

# Possible Service Scenarios



- **Full Support**
  - ITSD provides cradle-to-grave support service
  - Fully funded by overhead
  - High quality support
- **Cost Sharing**
  - ITSD provides cradle-to-grave support service
  - Participating divisions pay 25% of ongoing costs
  - High quality support
- **Start Up**
  - ITSD provides prepurchase consulting, setup and configuration
  - Ongoing systems admin performed by scientific division
  - ITSD will provide cybersecurity
  - Additional ITSD services on T&M basis
  - Cluster configuration and systems admin must meet certain criteria
  - Quality can be variable

# Overhead Cost for Program (\$K)

(Assume Jan. 1 Start<sup>1</sup>)



Year Option	FY 03	FY 04 <sup>2</sup>	FY05 <sup>2</sup>	3 yr Total
<b>Full Support</b>	<b>390</b>	<b>205</b>	<b>185</b>	<b>780</b>
<b>Cost Sharing</b> (Full Support with 25% recharge)	<b>360</b>	<b>135</b>	<b>110</b>	<b>605</b>
<b>Start Up<sup>3</sup></b>	<b>245</b>	<b>85</b>	<b>50</b>	<b>380</b>

<sup>1</sup> Cluster installations staged over the first year

<sup>2</sup> Assumes \$180K existing Scientific Computing OH in FY04 and FY05

<sup>3</sup> Divisions can pay T&M for support after initial start up

# Typical Costs to Projects (\$K)

(Assume Jan. 1 Start<sup>1</sup>)



		FY 03	FY 04	FY 05	Total 3 year Cost to Project
<b>Full Support</b>	<b>Cluster Purchase<sup>1</sup></b>	<b>75-125</b>	-	-	<b>75-125</b>
	<b>Ongoing Support</b>	-	-	-	-
	<b>Facilities</b>	-	-	-	-
<b>Cost Sharing</b> (25% Recharge of Ongoing Support)	<b>Cluster Purchase<sup>1</sup></b>	<b>75-125</b>	-	-	<b>75-125</b>
	<b>Ongoing Support<sup>2</sup></b>	<b>5</b>	<b>10</b>	<b>10</b>	<b>25</b>
	<b>Facilities</b>	-	-	-	-
<b>Start Up</b> (Division support 25% Postdoc)	<b>Cluster Purchase<sup>1</sup></b>	<b>75-125</b>	-	-	<b>75-125</b>
	<b>Ongoing Support<sup>2</sup></b>	<b>15</b>	<b>15</b>	<b>15</b>	<b>45</b>
	<b>Facilities</b>	-	-	-	-

<sup>1</sup> Cluster purchase cost will vary depending on size and configuration

<sup>2</sup> General rate burden of 45% applied where applicable

# Overall Costs to Lab for Scientific Computing Support (\$K)<sup>1</sup>



	<b>3 yr Cost To OH</b>
<b>Full Support</b>	<b>780</b>
<b>Cost Sharing (25% Recharge for Ongoing Support)</b>	<b>605</b>
<b>Start Up (Division support 25% Postdoc)</b>	<b>380</b>

<sup>1</sup> Does not include purchase price of clusters nor \$180K existing Scientific Computing OH

# Overall Costs to Lab for Scientific Computing Support (\$K)<sup>1</sup>



	<b>3 yr Cost To OH</b>	<b>3 yr Support Cost to Projects (10)</b>
<b>Full Support</b>	<b>780</b>	<b>0</b>
<b>Cost Sharing</b> (25% Recharge for Ongoing Support)	<b>605</b>	<b>250</b>
<b>Start Up</b> (Division support 25% Postdoc)	<b>380</b>	<b>450</b>

<sup>1</sup> Does not include purchase price of clusters nor \$180K existing Scientific Computing OH

# Overall Costs to Lab for Scientific Computing Support (\$K)<sup>1</sup>



	3 yr Cost To OH	3 yr Support Cost to Projects (10)	Total 3 yr Cost w/General Rate burden
<b>Full Support</b>	<b>780</b>	<b>0</b>	<b>780</b>
<b>Cost Sharing</b> (25% Recharge for Ongoing Support)	<b>605</b>	<b>250</b>	<b>855</b>
<b>Start Up</b> (Division support 25% Postdoc)	<b>380</b>	<b>450</b>	<b>830</b>

<sup>1</sup> Does not include purchase price of clusters nor \$180K existing Scientific Computing OH

# Overall Costs to Lab for Scientific Computing Support (\$K)<sup>1</sup>



	3 yr Cost To OH	3 yr Support Cost to Projects (10)	Total 3 yr Cost w/General Rate burden	Total 3 yr Cost w/o General Rate burden
<b>Full Support</b>	<b>780</b>	<b>0</b>	<b>780</b>	<b>780</b>
<b>Cost Sharing</b> (25% Recharge for Ongoing Support)	<b>605</b>	<b>250</b>	<b>855</b>	<b>780</b>
<b>Start Up</b> (Division support 25% Postdoc)	<b>380</b>	<b>450</b>	<b>830</b>	<b>680</b>

<sup>1</sup> Does not include purchase price of clusters nor \$180K existing Scientific Computing OH