



**SLAC**  
**SCIENTIFIC COMPUTING** June 20-21, 2011  
**workshop**  
Redwood Rooms, Research Office Building 4E



The **SLAC Scientific Computing Workshop** is a forum to learn about current activities and future directions in Scientific Computing across the Lab for the purpose of sharing ideas and information. Topics include **data management, algorithms, simulation, visualization, collaborative tools and emerging hardware architectures**. The intent of this workshop is to lead to new collaborative efforts.

We are soliciting 5-minute "lightning" talks with a deadline of May 27. The agenda, registration and talk sign-up can be found at:

<http://tinyurl.com/scw2011>

**SLAC**

## Scientific Computing Closeout

Amber Boehnlein,  
Richard Dubois  
*June 21, 2011*

Local Organising Committee:

Amedeo Perazzo  
Amber Boehnlein  
Arno Candel  
Ashley Deacon  
Richard Dubois  
Randy Melen  
Brian Moritz  
Tony Johnson  
Jacek Becla  
Travis Brooks



# Workshop Goals

---

- Find out what is going on across the Lab
  - Also from Stanford—ICME
- Getting acquainted
- Identify needs and roadblocks
- Identify skill sets, both in the directorates and in CD
  - What is already known in various parts of the Lab that could address overall needs
  - CD's ability to serve the Lab's projects
- First of an annual event

# DM 1

lcls	fermi	babar	lsst	jccg	xrd	xray		score
			2	2			automation	4
2							bbcp	2
	2		2				fits	4
2			2	2			gpus	6
			2				hdf5	2
2	2	2			2		hpss	8
	2		2				htm	4
		2	2	2			io optimizations	6
2			2				iRods	4
1	1	2	2	2			long term archive	8
2			2				lustre	4
						2	matlab	2
2			2				MySQL	4
	2			2			oracle	4
1	2	2	2				provenance	7

# DM2

---

Mentioned, Used, Needed

Automation, `bbcp`, `fits`, `gpus`, `hdf5`,  
`hpss`, `htm`, **io optimizations**,  
`iRods`, **long term archive**, `lustre`,  
`matlab`, `MySQL`, `Oracle`, **provenance**,  
`reDDnet`, `root`, **xrootd**, `xtc`

# Simulation

---

- Simulation is essential to all the SLAC scientific projects
- Central support of computational resources (mainly CPU, but storage as well)
- Simulation tools
  - PPA, Radiation Protection -- Geant4, Fluka, (MCNP)
  - Accelerator -- ACE3P, Impact, Elegant, Genesis, etc.
  - SSRL/LCLS -- ?
    - Cross-validation over tools and share the outcomes
    - Open forum to interested parties

# Algorithms Summary

---

- Beyond model fitting
  - Emphasizing instead reconstruction and discovery
  - Finding efficient algorithms for characterizing data
- Looking for patterns in large, multidimensional datasets
  - known unknowns:
    - CTs to identify cosmic rays vs photons, pulsars vs blazars
  - unknown unknowns:
    - unsupervised clustering in Sloan, stock market, biomedical data
    - reconstructing unknown source intensities from noisy data (EMC algorithm)
- Other themes
  - Mapping of problem domains (quantum mechanics, robotics)
  - Importance of visualization (visualization of DQ clustering; TMine GUI; protein folding)
  - Sharing Expertise (ICME; this workshop)

# Summary: Visualization and Emerging Architectures

---

GPUs can significantly speed up scientific computing in many fields

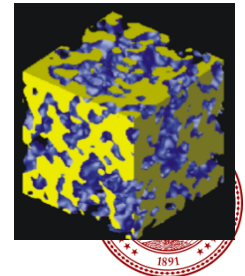
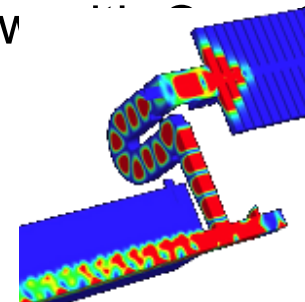
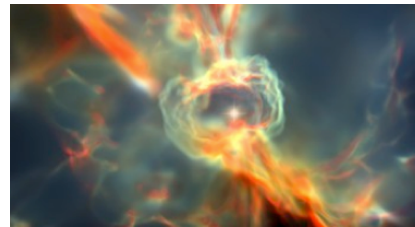
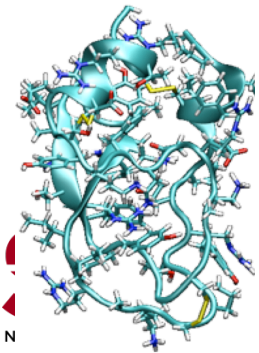
Prerequisite for optimal performance:

Problems

- fit well to stream processing (SIMD) model
- are massively parallel, have many lightweight threads
- have high computational load (high ratio of FLOPS vs Mem OPS)

Examples of successful GPU applications include:

- Quantum chemistry (CUDA)
- Fermi data analyses (CUDA)
- Visualization in numerical cosmology (OpenGL shading language)
- Visualization in accelerator rf design (ParaView, OpenGL)



# Next Steps

---

- Develop data management plans (DMPs)
- Assess ability of existing manpower and hardware resources to address the needs
- Assemble Strategy Plan based on DMPs, budgets etc
  - Tied in to Lab agendas
  
- We need YOU to participate!



# Feedback from the Workshop

---

- Tremendous amount going on at the Lab!
- Communications will be key
  - Entries already going into the forum!
  - How to encourage participation with the SCSC
    - Quarterly town halls?
    - Take Tony's suggestion for open meetings under advisement

# Processes we need

---

- modifying recharge formula
- buying from new vendors and novel hardware. Relationship to Lance's alternate vendor search?
- keeping CD informed of upcoming purchases, and setting priorities if needed to resolve bottlenecks.
- evaluating licences and deciding who pays
- CD to consult directorates on resource changes affecting scientific computing.
- getting SCSC input into the Operations Directorate Agenda and the Operations Directorate Strategic Business Plan
- Annual reviews of CD Scientific Computing and M&S expenditures
- technical design reviews
- for CD to consult about projects within the Unix/storage team.

# Working together

---

- [sci-comp@slac.stanford.edu](mailto:sci-comp@slac.stanford.edu) for announcements
- [forum.slac.stanford.edu](http://forum.slac.stanford.edu) to discuss issues
- Monthly scientific computing seminars to share knowledge in person
  
- SCSC as your representatives to Lab management
  - [scsc-l@slac.stanford.edu](mailto:scsc-l@slac.stanford.edu) to contact them
  - <https://confluence.slac.stanford.edu/display/scscpub/Scientific+Computing+Steering+Committee+Public>
  - Meeting agendas and minutes are posted:  
<https://confluence.slac.stanford.edu/display/scscpub/SCSC+Meeting+Agendas>
  - Put a “watch” on the agenda page to be notified of updates!