HEP and Non-HEP Computing at a Laboratory in Transition

Richard P. Mount

Director: Scientific Computing and Computing Services

Stanford Linear Accelerator Center

CHEP 2007 September 6, 2007

The Transition

- 1962: SLAC founded to construct and exploit a 2-mile linear accelerator;
- 1973: SSRL founded to exploit sychrotron radiation from SLAC's SPEAR storage ring;
- 2003: KIPAC (Kavli Institute for Particle Astrophysics and Cosmology) founded at SLAC/Stanford;
- 2005: LCLS (Linac Coherent Light Source) high energy X-ray laser construction approved;
- 2008: Last year of BaBar data taking. Last year of operation of SLAC accelerators for experimental HEP;
- 2009: BES (Photon Science funding) takes ownership of SLAC accelerators. BES becomes the largest funding source for SLAC. LCLS operation begins;
- The future:
  - Vigorous Photon Science program;
  - Vigorous Astrophysics and Cosmology program;
  - Potentially vigorous HEP program.

Cultural Evolution in SLAC Computing

- Experimental HEP:
  - Large, organized collaborations
  - Computing recognized as vital
  - Detailed planning for computing
  - Expectation of highly professional computing
- Astrophysics and Cosmology (Theory)
  - Individual PIs or small collaborations
  - Computing recognized as vital
  - Desire for agility
  - “Professional” approach viewed as ponderous and costly
- Astronomy
  - Increasingly large collaborations and HEP-like timescales
  - Computing recognized as vital
  - Detailed (8 years in advance of data) planning for computing
  - Expectation of highly professional computing
- Photon Science
  - Light sources are costly facilities
  - Photon science has been “small science” up to now (“turn up for 3 days and take away a DVD”)
  - No large, organized collaborations
  - Computing not considered a problem by most scientists
  - Detailed planning would be a waste of time
  - “Wait for the crisis and then we will get the resources to fix the problems”

Technical Evolution in SLAC Computing

- Experimental HEP:
  - Throughput-oriented data processing, trivial parallelism
  - Commodity CPU boxes
  - Ethernet switches ideal
  - Experiment-managed bulk disk space
  - Challenging data-management needs
- Astrophysics and Cosmology (Theory)
  - Visualization
  - Parallel Computing
  - Shared-memory SMP, plus
  - Infiniband MPI Clusters
  - Lustre and other “fancy” file systems
  - Macs and Linux boxes
- Astronomy
  - Visulaization
  - Challenging data management needs
  - Lustre and other “fancy” file systems
  - Macs and Linux boxes
- Photon Science
  - Computer science challenges (extract information from a million noisy images)
  - MPI clusters needed now
  - Significant needs for bandwidth to storage
  - Other needs unclear

DOE Scientific Computing Annual Funding at SLAC

- Particle and Particle Astrophysics
  - $14M SCCS
  - $5M Other
- Photon Science
  - $0M SCCS (ramping up real soon now)
  - $1M SSRL?
- Computer Science
  - $1.5M

SLAC Scientific Computing

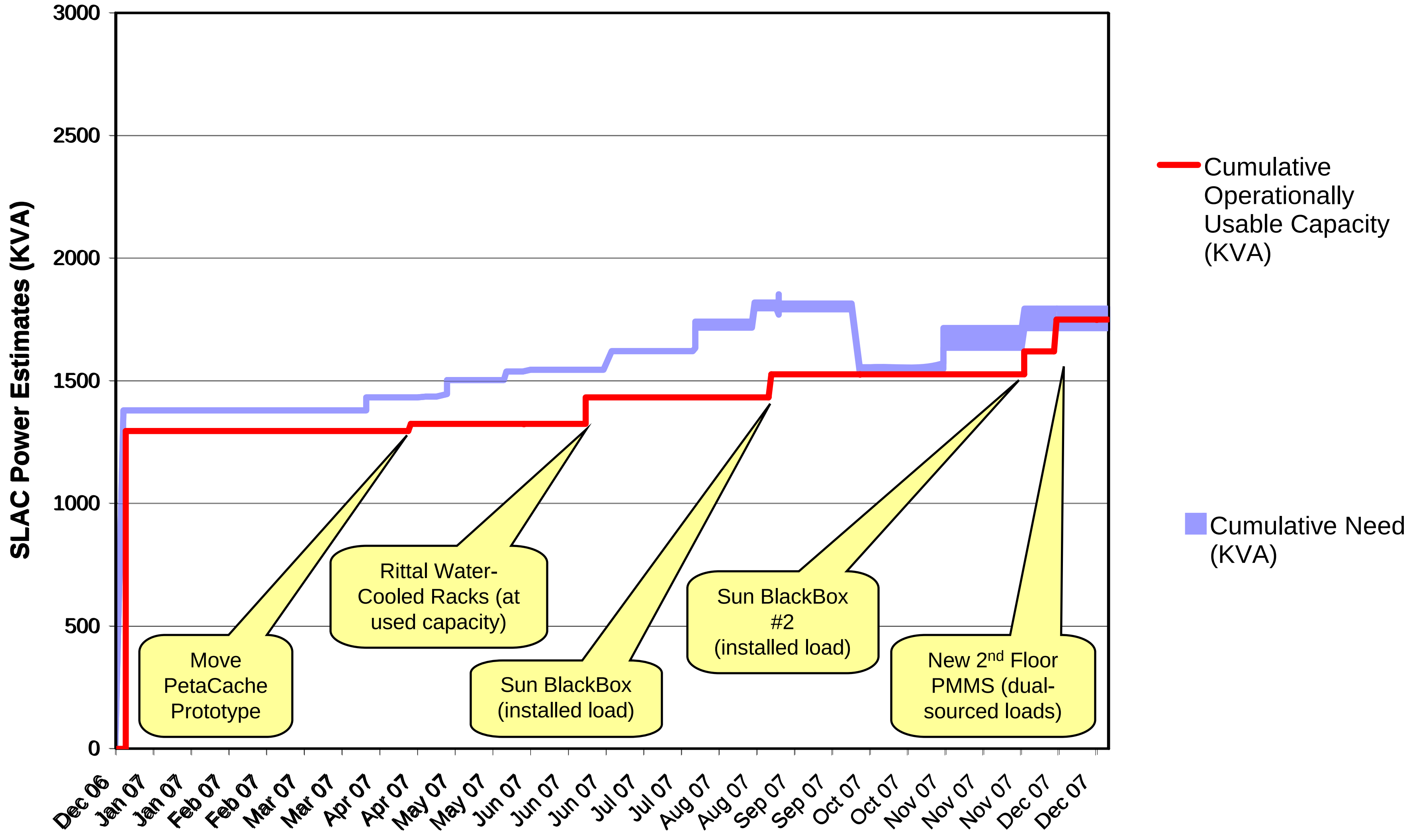
Science	Science Goals	Computing Techniques
BaBar Experiment (winds down 2009-2012)	Measure billions of collisions to understand matter-antimatter asymmetry (why matter exists today)	High-throughput data processing, trivially parallel computation, heavy use of disk and tape storage
Experimental HEP (LHC/ATLAS, ILC …)	Analyze petabytes of data to understand the origin of mass	High-throughput data processing. trivially parallel computation, heavy use of disk and tape storage
Accelerator Science	Current: World-leading simulation of electromagnetic structures; Future: Simulate accelerator (e.g. ILC) behavior before construction and during operation	Parallel computation, visual analysis of large data volumes
Particle Astrophysics (mainly simulation)	Star formation in the early universe, colliding black holes, …	Parallel computation (SMP and cluster), visual analysis of growing volumes of data
Particle Astrophysics Major Projects (GLAST, LSST …)	Analyze terabytes to petabytes of data to understand the dark matter and dark energy riddles	High-throughput data processing, very large databases, visualization
Photon Science (biology, physics, chemistry, environment, medicine)	Current: SSRL – state-of-the art synchrotron radiation lab; Future LCLS – Femtosecond x-ray pulses, “ultrafast” science, structure of individual molecules …	Medium-scale data analysis and simulation. Visualization. High throughput data analysis and large-scale simulation. Advanced visualization
PetaCache – New Architectures for SLAC Science	Radical new approaches to computing for Stanford-SLAC data-intensive science	Current focus: massive solid-state storage for high-throughput, low-latency data analysis

Power and Cooling

- Near future: difficult
- Medium term: very difficult
- Long term: totally scary

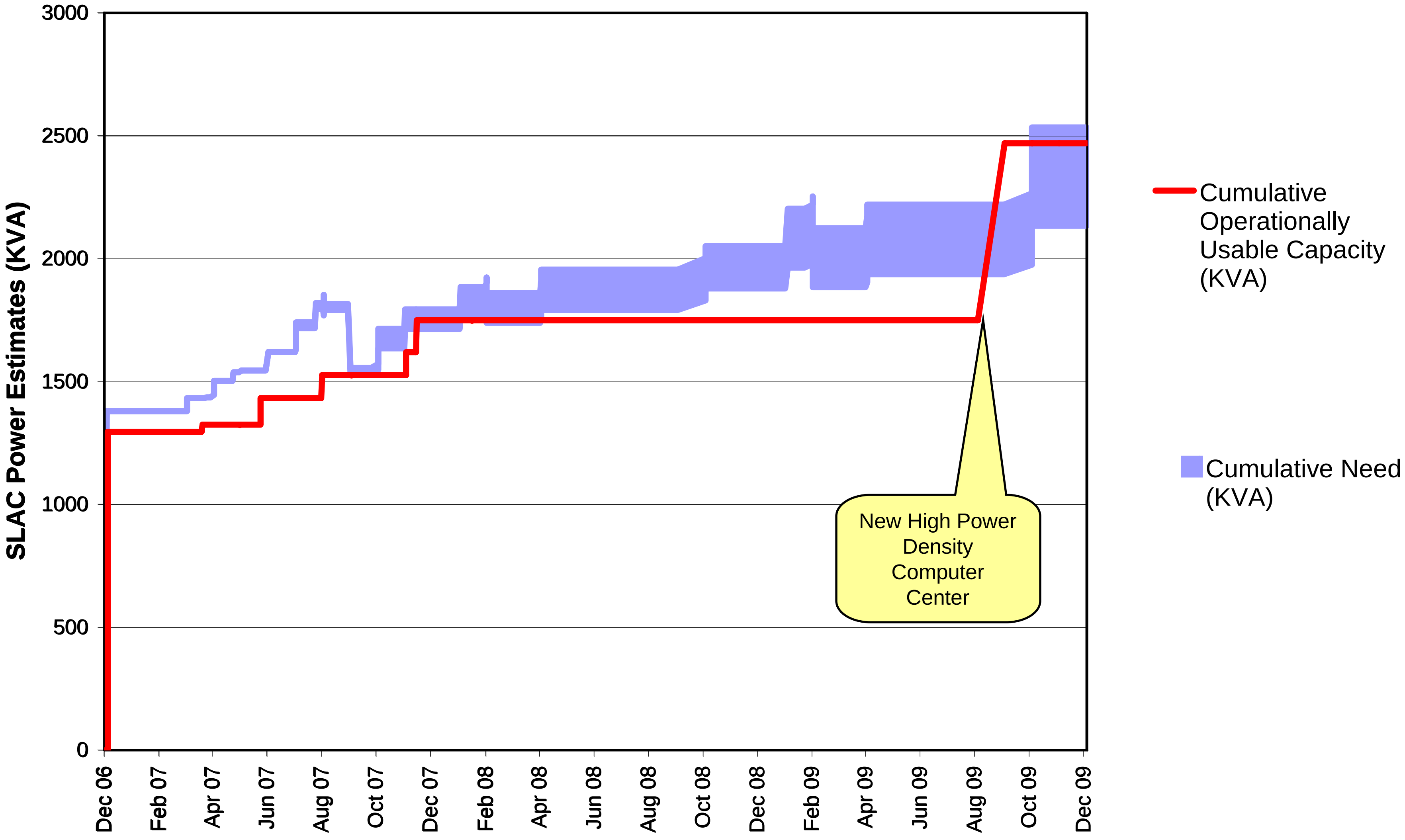
Near Future

SLAC Computing infrastructure (Need vs. Capacity)



Medium Term

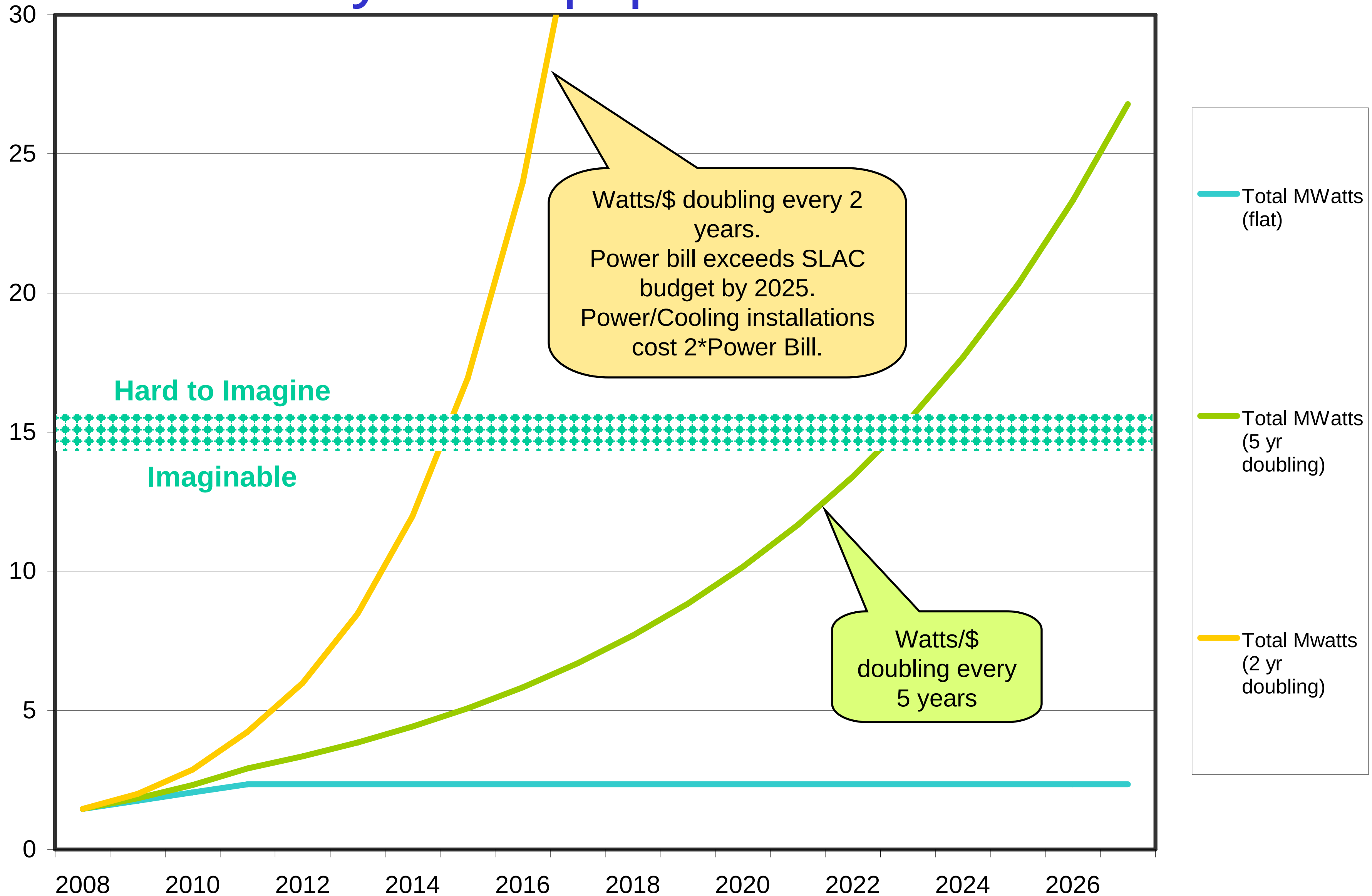
SLAC Computing infrastructure (Need vs. Capacity)



And the Longer-Term Future?

- Six major programs each justifying $Millions/year of hardware
- Assume modest level of funding success: $6M/years
- Current power consumed:
  - CPU: 0.08 Watts/$
  - Disk: 0.05 Watts/$
  - Cooling systems: ~50% on top of these
- Future evolution assumptions:
  1. Watts/$ is flat
  2. Watts/$ doubles every 5 years (consistent with recent experience)
  3. Watts/$ doubles ever 2 years (à la James Sexton)

SLAC Computing MWatts
$6M/year Equipment Purchases



Short Term Fixes for Power and Cooling

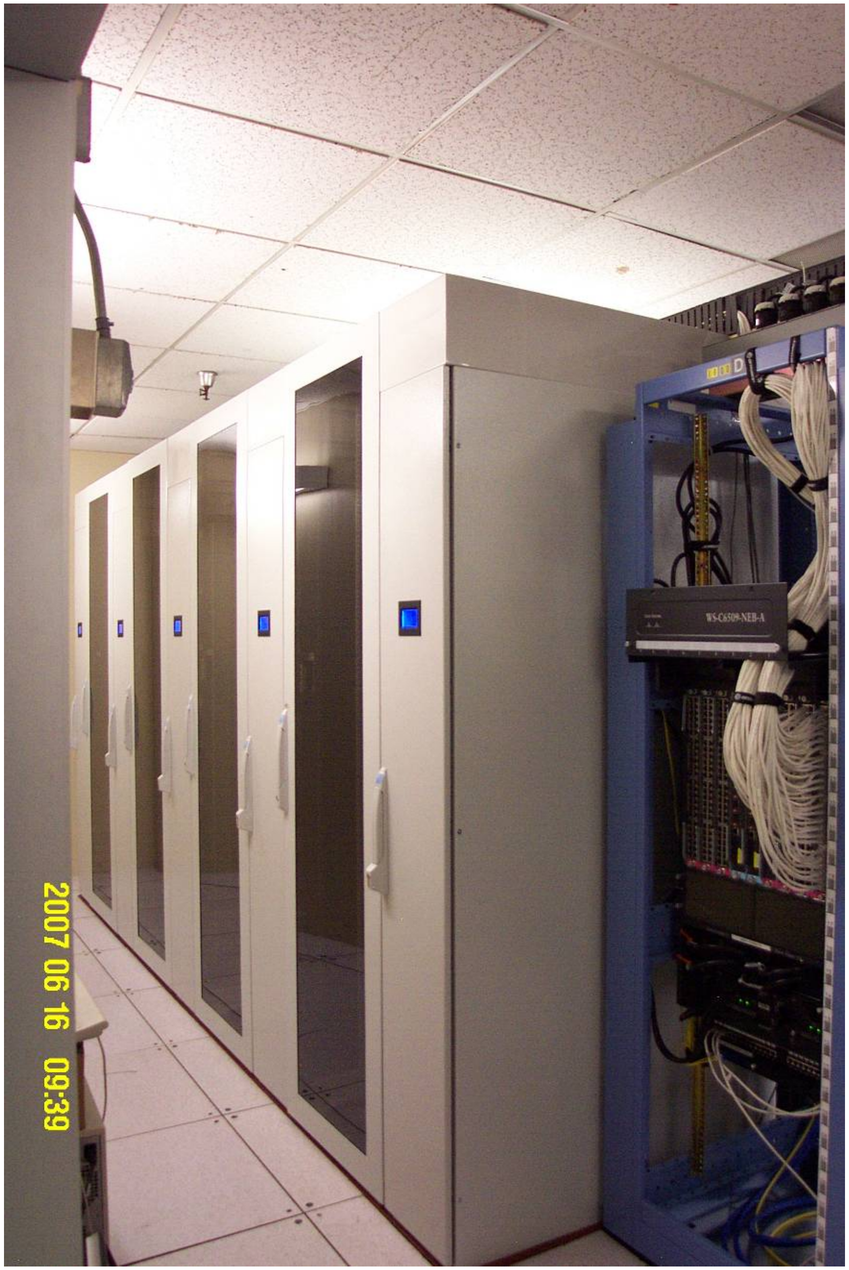
- Rittal Water-Cooled Racks
  - Eight 40RU racks, up to ~25KW per rack
  - Online end June

- Sun BlackBox
  - Holds 252 1RU servers plus networking
  - Up to 200KW payload possible
  - Online September 17

- Replace “old?” “junk?”
  - Old can mean > 2 years old
  - Junk can mean P4-based machines
  - Sun “Thumpers” are 4 to 5 times as power efficient per TB as our 2-year old sun disks + servers.

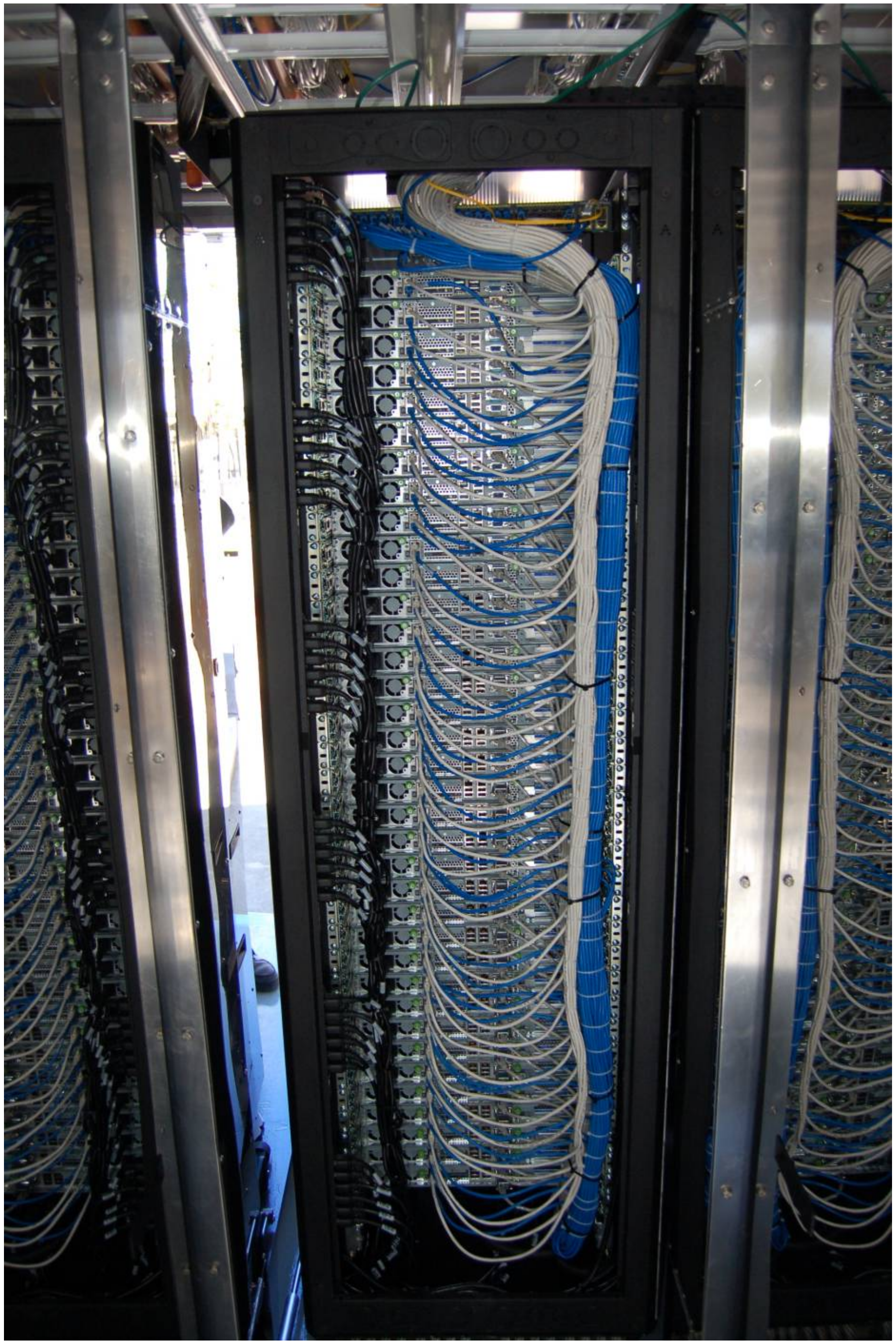
Rittal Water-Cooled Racks

- 30cm wide heat exchanger racks between each equipment rack
- 20KW maximum per equipment rack
- 8 racks installed (photograph shows four).
- Contents, CPU for:
  - ATLAS Tier 2 (June 2007)
  - GLAST (June 2007)
  - BaBar, Noma/Tori replacement (September 2007)
  - ATLAS Tier 2 (end 2007)




Sun
microsystems


Sun
microsystems



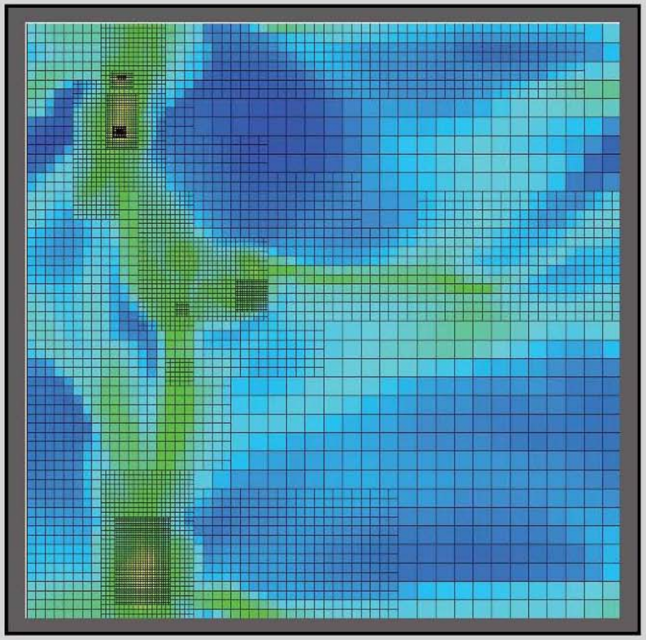
New Science Directions

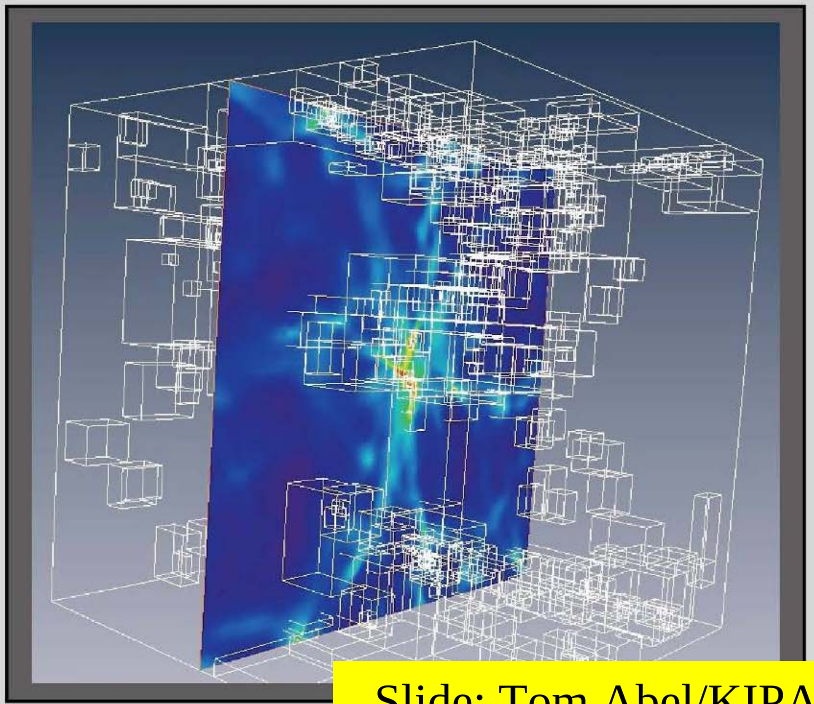
Astronomy, Astrophysics and Cosmology

Slide: Tom Abel/KIPAC

Cosmological Adaptive Mesh Refinement

- **Enzo:** Bryan and Norman 1997-; Abel et al 97; Anninos et al 97; Bryan, Abel & Norman 2002; O'Shea et al; Abel, Wise & Bryan 2006
  - 87,000 lines of code in C++ and F77
  - Cosmological Radiation Hydrodynamics adapting in space and time
  - Dynamic range up to 1e15 using quadruple precision coordinates in space and time
  - Dynamically load balanced parallel with MPI
  - Gravity, DM, Gas, Chemistry, Radiation, star formation & feedback
  - Current new Developments at KIPAC: exact 3D radiation transport, very high density chemistry, HD & fine structure line cooling, relativistic hydro, passive MHD, new visualization toolkit





Slide: Tom Abel/KIPAC

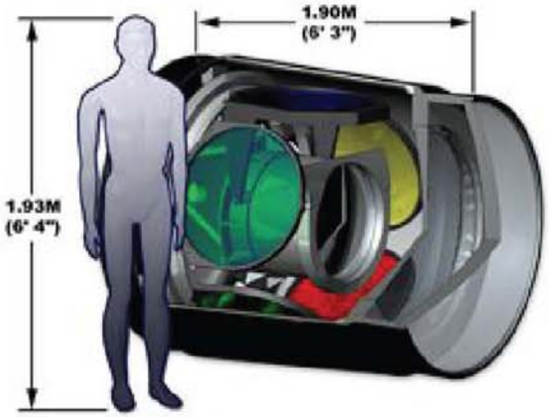
Summary

- First Stars are isolated and very massive: I.e. form *before* Galaxies
- Enormous impact on subsequent structure
  - Heavy Elements, Entropy, Turbulence
- Bleeding edge of computational astrophysics, 1e8 larger dynamic range than anybody else.
  - completely new algorithms (RT, pMHD, relativistic AMR, data paradigm), 128 bit arithmetic (for coordinates), parallelization and dynamic load balancing, custom visualization tools, PIC plasma applications very similar to cosmological N-body ...
- Looking forward to learn with Photon Sciences about the LCLS algorithmic

Slide: Tom Abel/KIPAC

Strong lensing with LSST

Large Synoptic Survey Telescope



High etendue survey telescope

- 6m effective aperture
- 10 sq degree field
- 24.5 mag in 30 seconds
- Visible sky mapped in three nights
- Ten year *movie* of the sky

Slide: Phil Marshall/UCSB

New Science Directions

Photon Science

LCLS: Linac Coherent Light Source

LUSI: LCLS Ultrafast Science Instruments

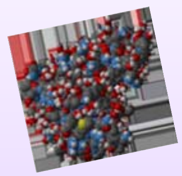
Photon Science at LCLS
Opportunities on Several Fronts

- LCLS will begin operations in 2009
- LCLS DAQ will have LHC-like rates to storage:
  - Design of the complete DAQ systems is just starting
- LCLS offline computing has many "opportunities":
  - Opportunity to define where offline analysis will be done
  - Opportunity to define the scale of offline analysis (e.g. will Coherent Crystal Imaging need 40,000 cores?)
  - Opportunity to create science/computer science teams to write algorithms and framework(s)
  - Opportunity to acquire funding for these activities
- Opportunity to make a small science/big science transition successful, and fun

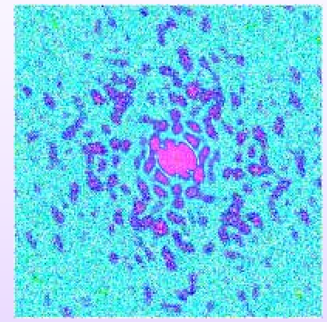
Coherent Imaging of Single Molecules

- *Diffraction from a single molecule:*

single LCLS pulse



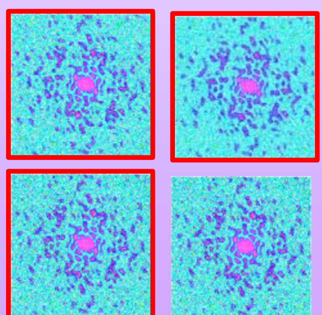
unknown orientation



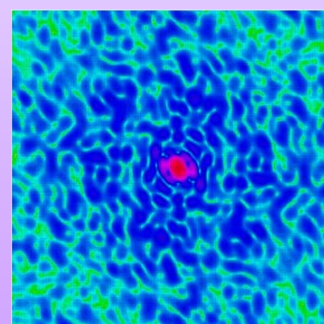
noisy diffraction pattern of unknown orientation

- *Combine 10^5 to 10^7 measurements into 3D dataset:*

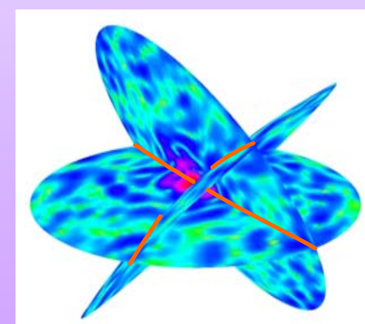
Classify/sort



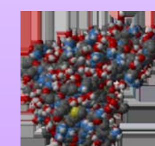
Average



Alignment



Reconstruct by Oversampling phase retrieval



The highest achievable resolution is limited by the ability to group patterns of similar orientation

Gösta Huldt, Abraham Szöke, Janos Hajdu (J.Struct Biol, 2003 02-ERD-047)

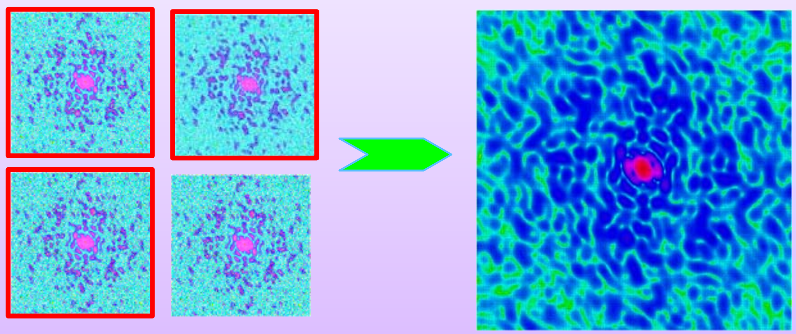
Miao, Hodgson, Sayre, PNAS 98 (2001)

Classification/Averaging

- Combine 10^5 to 10^7 measurements into 3D dataset:
  with15%/25% efficiency, 10^7 → 40 days
  with 30%/50% efficiency, 10^7 → 7 days
  → Long duration adds additional complications

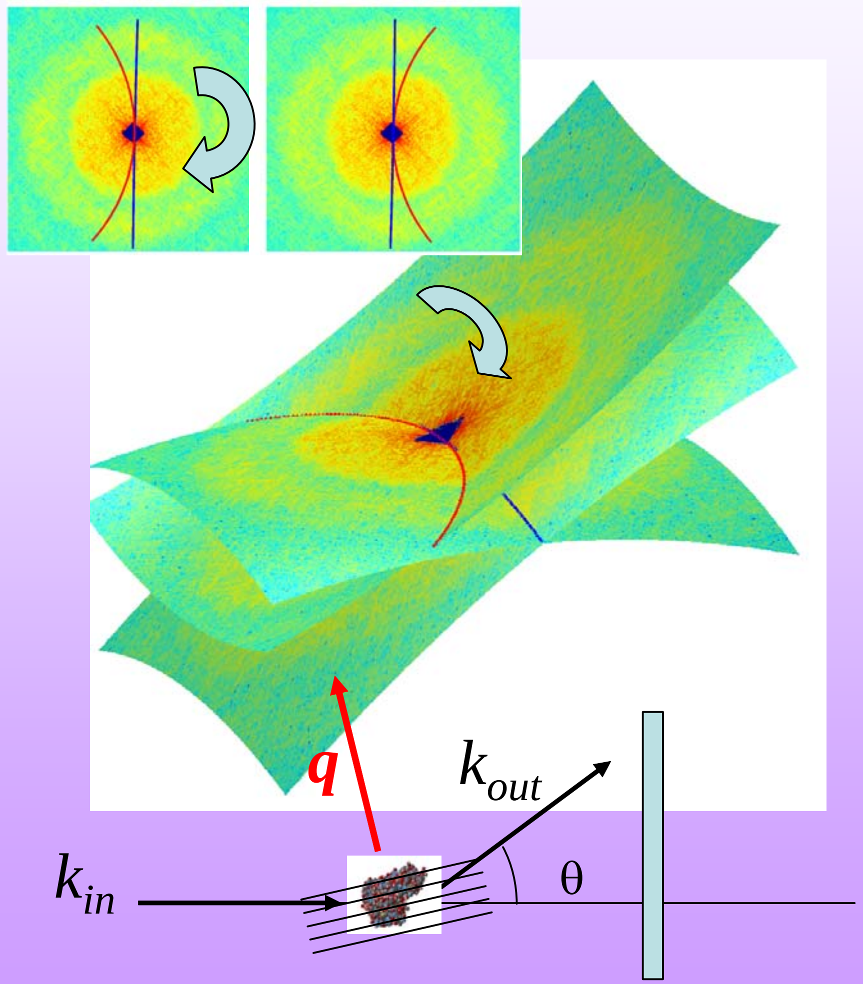
Classify/sort

Average

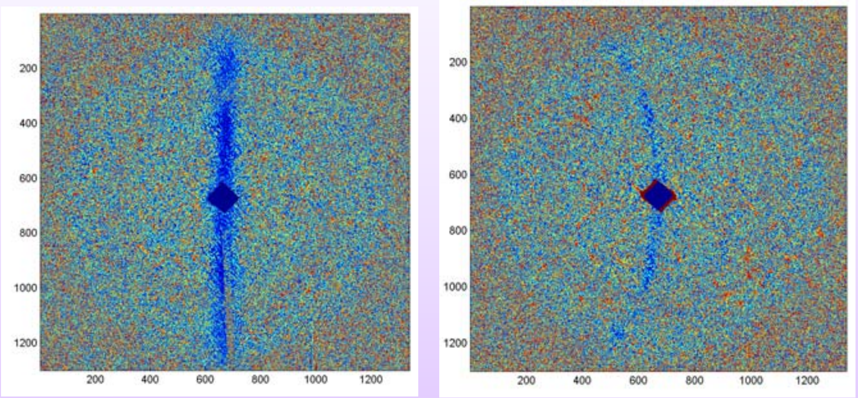


- Normalize frames properly
  intensity, positional, energy shifts
- Establish many (~1000) dissimilar templates
- Compare each frame to templates to find match
  rotate each frame about beam axis in increment
- Average similar frames
  → 10^{17} FLOPS

Computational Alignment



Experimental Data (ALS)

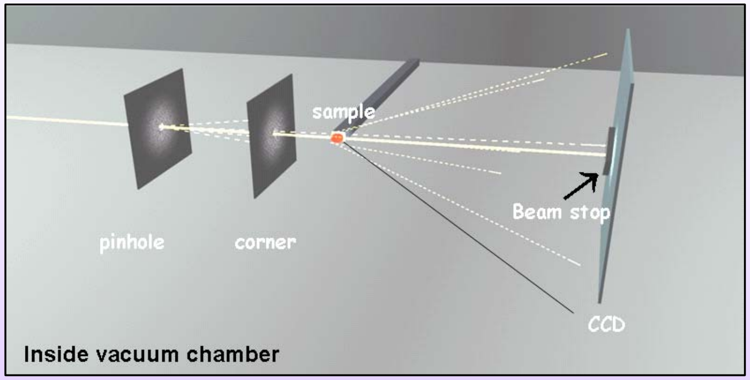


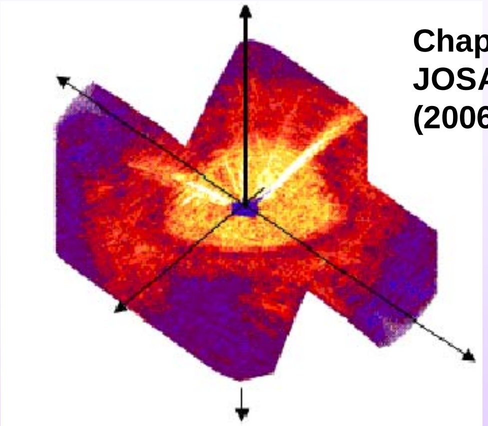
Difference of pyramid diffraction patterns 10° apart, Gösta Huldt, U. Uppsala

"The number currently used to obtain high-resolution structures of specimens prepared as 2D crystals, is estimated to require at least 10^{17} floating-point operations"
R. M. Glaeser, J. Struct. Bio. 128, (1999)

"Computational Alignment" requires large computational power that can only be provided by performing offline analysis
➔ Save first, and Analyze later

Computational Diffraction Imaging





Chapman et al. JOSAA 23, 1179 (2006)

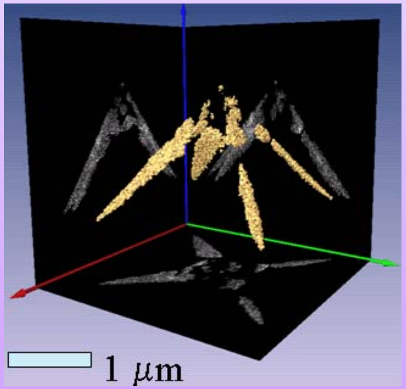
Optimization

- 10^9 nonlinear equations
- 10^9 unknowns

Shrinkwrap/HIO algorithm, S. Marchesini et al, PRB 68, 140101(R) (2003)

3D size	Memory (GB)	Time (Hours)
256^3	0.6	0.2
512^3	4.7	1.5
1024^3	38	14
2048^3	304	??
2000 iterations, 2FFTs each 32 CPU G5 Infiniband cluster		

Crandall, et al. Advanced Comp. Group, Apple Computer (2004).



Longer Term Trend

For CXI Detector: pixel number x2/3 years			
	Average (in 2009)	2012	2015
Peak Rate (Gigabit/s)	1.95	3.90	7.80
Success Rate (%)	10%	30%	50%
Ave. Rate (Gigabit/s)	0.20	1.17	3.90
Daily Duty Cycle (%)	25%	50%	75%
Accumulation (Terabyte/day)	0.53	6.31	31.6
Yearly Uptime (%)	25%	75%	90%
Accumulation (Petabyte/year)	0.048	1.73	10.4

Pixels will go up, machine will be more stable and measurement techniques more refined

In Conclusion

- What's exciting about scientific computing at SLAC?
  - The new adventure of the LCLS program with its many computing challenges
  - The new adventure of Particle Astrophysics and Cosmology with unbounded appetite for computing
  - The old adventure of high energy physics experimentation with constant new challenges
  - The dream of fully simulating accelerators
  - Allowing this excitement to inspire new developments and wild ideas in computing that will greatly increase the success of the science program.