LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-218445

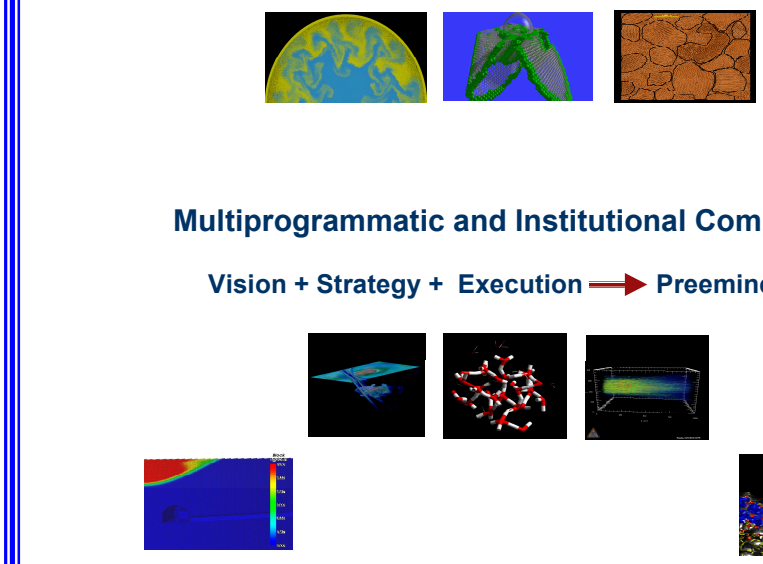
2005 White Paper on Institutional Capability Computing Requirements

B. Carnes, M. McCoy, M. Seager

January 26, 2006

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

2005 White Paper on Institutional Capability Computing Requirements



Prepared by
Brian R. Carnes
Michel G. McCoy
Mark K. Seager

Integrated Computing and Communications Department
Computation Directorate
Lawrence Livermore National Laboratory

and the open computing scientists at LLNL

for the
Institutional Computing Executive Group
Associate Director for Computation
Deputy Director for Science and Technology

July 27, 2005

Contents

Executive Summary

This paper documents the need for a significant increase in the computing infrastructure provided to scientists working in the unclassified domains at Lawrence Livermore National Laboratory (LLNL). This need could be viewed as the next step in a broad strategy outlined in the January 2002 White Paper (UCRL-ID-147449) that bears essentially the same name as this document. Therein we wrote: "This proposed increase could be viewed as a step in a broader strategy linking hardware evolution to applications development that would take LLNL unclassified computational science to a position of distinction if not preeminence by 2006." This position of distinction has certainly been achieved. This paper provides a strategy for sustaining this success but will diverge from its 2002 predecessor in that it will:

- Amplify the scientific and external success LLNL has enjoyed because of the investments made in 2002 (MCR, 11 TF) and 2004 (Thunder, 23 TF).
- Describe in detail the nature of additional investments that are important to meet both the institutional objectives of advanced capability for breakthrough science and the scientists' clearly stated request for adequate capacity and more rapid access to moderate-sized resources.
- Put these requirements in the context of an overall strategy for simulation science and external collaboration.

While our strategy for Multiprogrammatic and Institutional Computing (M&IC) has worked well, three challenges must be addressed to assure and enhance our position. The first is that while we now have over 50 important classified and unclassified simulation codes available for use by our computational scientists, we find ourselves coping with high demand for access and long queue wait times. This point was driven home in the 2005 Institutional Computing Executive Group (ICEG) "Report Card" to the Deputy Director for Science and Technology (DDST) Office and Computation Directorate management.

> *"The ICEG recommends that M&IC acquire several 10 TFLOP/s class systems over the next 2–3 years to relieve the current serious shortage of mid-range capability computing. If M&IC does not acquire new systems soon, it will be hard for unclassified computing at LLNL to remain at the forefront of simulation science."*
>
> *"M&IC and LLNL's computational scientists have, in some sense, been victims of their own success. There are now many codes that can productively use thousands of processors for many days on a single run. A number of projects that can perform very high quality simulations did not receive time on Thunder. The MCR cluster currently offers less than 50% of the computing power needed to keep these projects moving forward at a reasonable rate, and the gap will rapidly increase if M&IC does not purchase new systems."*

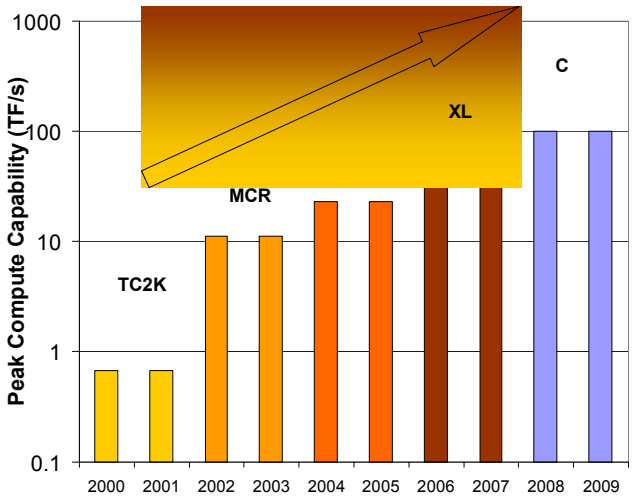
The second challenge is related to the balance that should be maintained in the simulation environment. With the advent of Thunder, the institution directed a change in course from past practice. Instead of making Thunder available to the large body of scientists, as was MCR, and effectively using it as a *capacity* system, the intent was to make it available to perhaps ten projects so that these teams could run very aggressive problems for breakthrough science. This

usage model established Thunder as a *capability* system. The challenge this strategy raises is that the majority of scientists have not seen an improvement in capacity computing resources since MCR, thus creating significant tension in the system. The question then is: "How do we address the institution's desire to maintain the potential for breakthrough science and also meet the legitimate requests from the ICEG to achieve balance?" Both the capability and the capacity environments must be addressed through this one procurement.[1]

The third challenge is to reach out more aggressively to the national science community to encourage access to LLNL resources as part of a strategy for sharpening our science through collaboration. Related to this, LLNL has been unable in the past to provide access for sensitive foreign nationals (SFNs) to the Livermore Computing (LC) unclassified "yellow" network. Identifying some mechanism for data sharing between LLNL computational scientists and SFNs would be a first practical step in fostering cooperative, collaborative relationships with an important and growing sector of the American science community.

We suggest three related steps as part of a broad strategy to sustain LLNL unclassified computing. These steps address the challenges described above in order.

1. **M&IC requests support for a procurement to put additional assets on the floor to carry the institution through FY06 at the 80 TF/s level.** We propose to procure an additional 45 TF for the institutional computing community. This augmentation would join Thunder (23 TF) and MCR (11 TF) for a total of ~80 TF peak capacity. The strategy is based on new 8-way SMP nodes that could be clustered into 5 TF scalable units (SUs). SUs can be aggregated arbitrarily to create more capable systems.



A possible platform path for institutional preeminence.

2. **M&IC will deliver capacity and capability at a 1:1 ratio of TFs, preserving balance while preserving potential for breakthrough simulation.** We propose to create a 40 TF capability system (with eight of the SUs) and then convert Thunder and the ninth proposed SU into capacity systems. M&IC would then have Thunder plus MCR and the stand-alone SU (5 TF) for 39 TF peak *capacity* along with 40 TF of new peak *capability*.

[1] National developments can provide useful ancillary data. Sandia has announced procurement of 59 TF/s of Dell–Intel hardware as an institutional investment. This system is balanced to achieve affordable cost. It is clear that other institutions have resonated with the importance of enabling their technical users at scale and are confident in the funding models they have developed. While we are not requesting funding for ~60 TF here to match that peak, we do have interest in maintaining the computing infrastructure our institution needs to be a world leader in simulation science.

This increases M&IC capacity by 28 TF to meet the requests from the ICEG for additional capacity. The institutional interest in a machine for breakthrough science will be met by the new 40 TF system, almost doubling existing capability.[2]

3. **Access to the institutional computing cycles will be achieved via proposal processes. There will be a mechanism for outreach to external users who will collaborate with LLNL researchers**. Currently, M&IC runs two different proposal processes for the institution. The first, which we can call **LLNL S&T** efforts, is and would continue to be the main mechanism of institutional capacity access for LDRD and other researchers. The second we can call **LLNL Grand Challenge** efforts, and these represent LLNL researchers requesting large banks of computing cycles on the capability resource for breakthrough calculations.[3] We propose a third group, which we will call **National Security Challenge** efforts. Here an external and an internal principal investigator (PI) collaborate to write a proposal for access based on certain criteria, such as alignment with LLNL S&T thrust areas (like Aurora). Responding to the ICEG request to address the issue of access by SFNs, we are procuring a disk farm ("collaborative data oasis") to be placed on the "green" network. SFNs and other collaborators can log into this network and access data transferred there by institutional users. This will permit scientists to analyze the same data and will enable external collaboration. The three options for access to institutional computing cycles are outlined below.

    - **LLNL S&T Efforts**
        - The main mechanism of institutional access for LDRD and other researchers.
        - Strengthening this area is what the ICEG requested.
    - **LLNL Grand Challenge Efforts**
        - Can include external collaborators (but not necessary).
        - Similar to our current tier 1 and 2 Thunder efforts.
    - **National Security Challenge Efforts**
        - Might be termed "building on external strengths."
        - One external and one internal PI.
        - Adhere to strategic S&T thrust areas.
        - Show the external team is first rate and has a track record.
        - Reasonable prospect of being breeding nuclei of future internal research groups at LLNL.

These three steps will address all of the major issues raised by the ICEG and will accommodate the desire of the institution to maintain first-rate capability while achieving robust outreach to the external community.

The remainder of this document is focused on the details of successes, requirements, costs, and history of M&IC.

---

[2] The long-term plan would be to transition this new 40 TF into a capacity system in 2008 and replace it with a 100-TF-class computer for the institution. At that time, MCR would be retired.

[3] Continues the precedent established with Thunder.

Section 1 provides impacts and scientific results, highlighting the value of past institutional and programmatic investments in M&IC.

Section 2 addresses the customer requirements at a high level.

Section 3 provides some particulars of the system architecture that is planned for the SUs as well as the green network collaborative data oasis.

Section 4 lays out details of the expected approximate cost of the system and additional costs such as support infrastructure (including electrical power). Our requests for FY06, FY07, and FY08 are included. The collaborative data oasis will be procured with $850K in FY05 IGPE Capital Funds. Section 4 also includes a timeline for procurement and use of the system.

Section 5 gives some additional history of M&IC, including some detail on why it has been successful to date as an institutional effort.

Section 6 is the summary.

Appendix 1 presents much broader impacts and results as described by the scientists.

Appendix 2 provides detailed spreadsheets of requirements as gathered through survey.

1. Impacts and Results

1.1 Impact on S&T Base and Laboratory Programs

Computation is now a mainstream method in theoretical science at LLNL. Computation is essential at the level at which analytically intractable models are explored or complex multiphysics phenomena must be understood quantitatively. As we understand more and more truly basic science, the Laboratory is looking to computation to make the vital quantitative connections among disparate phenomena that constitute the foundation of both pure and applied science. The total computing needs of these projects far exceed the current M&IC capability and capacity.

> *"The large computational resources that are provided by M&IC for unclassified computing are a critical factor in the success of the high-pressure research that is being carried out in the Quantum Simulations Group. In particular, the past investments that the Laboratory has made in M&IC resources such as MCR and Thunder have greatly expanded the range of systems that we can investigate within a first-principles context. However, as more users are taking advantage of these resources, our research is currently being severely limited by long turnaround times for batch jobs."*
> *— Eric Schwegler, PAT Directorate, H Division*

> *"As a computational chemistry group, it is absolutely necessary for us to have these resources in order to continue making an impact in this large, competitive community. It should be pointed out that we have forged collaborations with scientific leaders based on LLNL's state-of-the-art hardware, software, and computational expertise. Our many recent successes are giving LLNL a unique opportunity to become a world leader in high-performance computational chemistry and materials science."*
> *— Larry Fried, Scientific Capability Leader, Chemistry under Extreme Conditions*

M&IC has worked closely with the Stockpile Stewardship Program (SSP), and through mutual leverage they have made LLNL a premier institution for computational and simulation science. Such standing is vital to the continued success of LLNL science programs and to the recruitment and retention of top scientists. M&IC constantly works at balancing investments to meet the widespread demand for capacity computing and provide cost-effective capability platforms. In fielding computing platforms and advanced data assessment capabilities, M&IC leverages the investments of the SSP for the mutual benefit of the SSP, other programs, and LLNL.

> *"For the past several years, the unclassified computing resources provided by M&IC and its predecessors have been essential in meeting our programmatic goals. PF3d modeling of NOVA and NEL experiments (hohlraums and CO_2 gasbags) accomplished using M&IC resources have contributed to a number of publications. Maintaining a cutting-edge computing environment in the unclassified arena will remain essential to us in continuing to meet our programmatic goals for several reasons. Interaction with experimentalists occurs on the unclassified networks (where their data and analysis reside), reporting physics results to our sponsors and the physics community is greatly facilitated by having our simulation results on the unclassified systems, and collaborations with our off-site scientific colleagues can only occur on the unclassified network."*
> *— Bert Still, Plasma Physics Group, AX Division, DNT*

Through strong and consistent LLNL investments, M&IC has grown into a powerful unclassified computing resource that is being used across LLNL to push the limits of computing and its application to simulation science. Through these efforts, LLNL has become a premier laboratory in simulation science, and this has resulted in world-class scientific insight and has facilitated the recruitment and retention of leading physical and computer scientists. All LLNL programs and projects are bolstered through these efforts. The Thunder and MCR clusters provide LLNL scientists with access to world-class unclassified systems. The Grand Challenge results included in Appendix 1 and the extensive M&IC Bibliography published separately are testament to the outstanding scientists and their groundbreaking work.

> *"Access and use of the M&IC computational resources has been essential and critical to our success. Atmospheric chemistry is extremely CPU intensive. Many groups have had to cut back on modeling, run at reduced chemistry and physics, and other methods to get simulation throughput. Access to M&IC has allowed us to move forward with quality models and carry out innovative and first-ever simulations."*
> *— Philip Duffy, Deputy Division Leader, Atmospheric Science Division*

1.2 Overview of Publications

M&IC has enabled a significant number of publications, presentations, and external collaborations for our LLNL scientists. The complete M&IC bibliography, available as a separate document, includes journal articles (282), presentations (95), other publications (32), and ongoing collaborations related to terascale resources. To provide a sense of the national impact and quality of this work, we list below nine of the more notable journal articles from the bibliography.

1. Ogitsu, T., F. Gygi, and G. Galli. "Imperfect Crystal and Unusual Semiconductor: Boron, a Frustrated Element." (submitted to *Nature*)
2. Gee, R.H., N. Lačević, and L.E. Fried. "Atomistic Simulation of Spinodal-assisted Polymer Crystallization." (submitted to *Nature Materials*)
3. Reed, E.J., M. Soljačić, R.H. Gee, and J.D. Joannopoulos. "Coherent Optical Photons from Shock Waves in Crystals." (in peer review for *Nature*)
4. Bonev, S.A., E. Schwegler, T. Ogitsu, and G. Galli. "A Quantum Fluid of Metallic Hydrogen Suggested by First-Principles Calculations." *Nature* **431** (2004): 669–672. [This article was the cover feature.]
5. Marcy, T.P., D.W. Fahey, R.S. Gao, P.J. Popp, E.C. Richard, T.L. Thompson, K.H. Rosenlof, E.A. Ray, R.J. Salawitch, C.S. Atherton, D.J. Bergmann, B.A. Ridley, A.J. Weinheimer, M. Loewenstein, E.M. Weinstock, and M.J. Mahoney. "Quantifying Stratospheric Ozone in the Upper Troposphere Using In Situ Measurements of HCl." *Science* **304**, no. 5668 (2004): 261–265.
6. Kuo, I-F.W. and C.J. Mundy. "An Ab Initio Molecular Dynamics Study of the Aqueous Liquid-Vapor Interface." *Science* **303**, no. 5658 (2004): 658–660.
7. Raty, J-Y. and G. Galli. "Ultradispersivity of Diamond at the Nanoscale." *Nature Materials* **2**, no. 12 (2003): 792–795.

8. Rau, G.H. and K. Caldeira. "Minimizing Effects of CO_2 Storage in Oceans." *Science* **295**, no. 5553 (2002): 275–276 (in Letters).

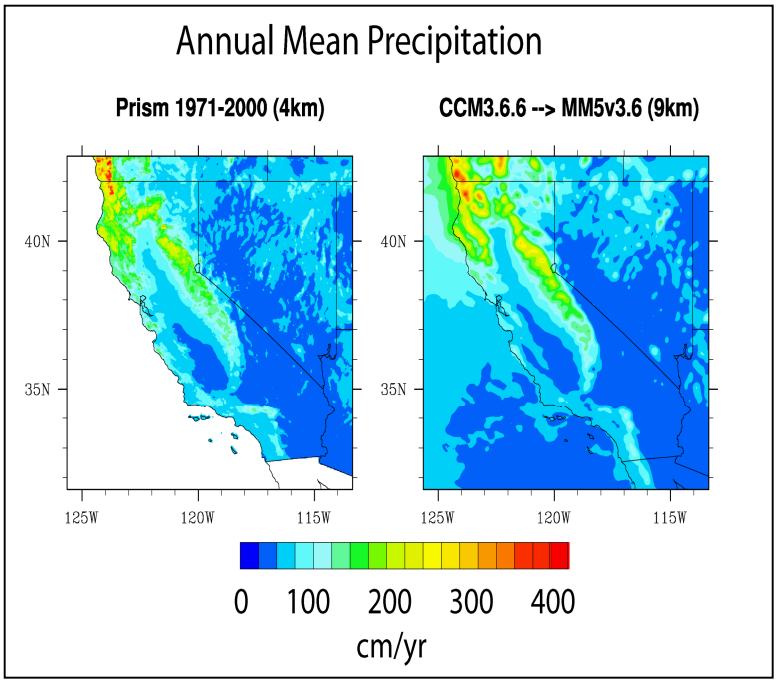
9. Caldeira, K. and P.B. Duffy. "The Role of the Southern Ocean in Uptake and Storage of Anthropogenic Carbon Dioxide." *Science* **287**, no. 5453 (2000): 620–622.

1.3 Selected Early Results

To provide another view of the impact of M&IC, we include below some detail on four projects featuring extremely compute-intensive work with some description, in the scientists' words, of why this work is valuable and of such interest to them. A more comprehensive overview of the institutional impact is provided by Appendix 1, which contains a much broader set of studies and results as described by the scientists as part of the requirements survey conducted for this White Paper.

1.3.1 High-Resolution Nested-Model Simulations of Annual Mean Precipitation

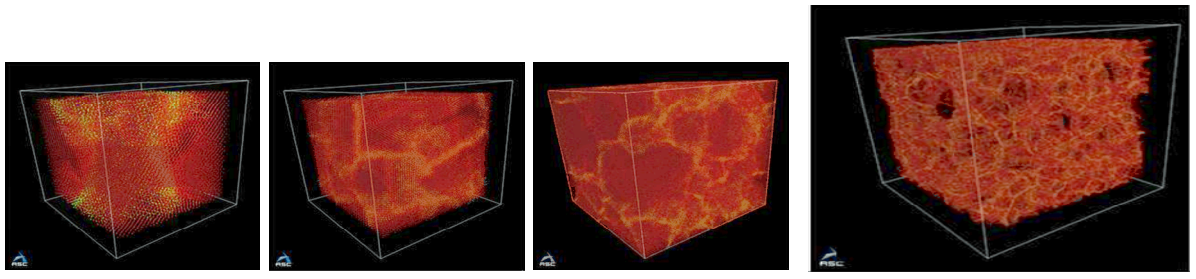
The figure shows annual mean precipitation in California–Nevada simulated by a regional (limited-domain) climate model at 9 km resolution nested within a global climate model at 75 km resolution. This nesting approach is widely used, but the grid dimensions here are about four times smaller than typical. A comparison to an observation-based estimate of precipitation (left panel) shows that the simulated precipitation has realistic spatial structure and minimal large-scale biases. The realistic spatial detail results from the high resolution of the nested model, which allows realistic representation of topographic variations. The minimal large-scale biases result from the high resolution of the driving global model.
— Phil Duffy



1.3.2 Molecular Dynamics Simulations of Pressure-Induced Resolidification in Tantalum

The short-range atomic structure and long-range grain structure created in a metal through rapid resolidification at high pressure and temperature are unknown, as is the timescale for the liquid-to-solid transition. Understanding the kinetics of resolidification and the morphology of the resolidified solid is vital to the accurate multiscale modeling of the mechanical properties of materials subject to extreme dynamic loading. We are performing molecular dynamics (MD) simulations to model the rapid resolidification of complex transition and actinide metals using advanced quantum-based many-body MGPT potentials that accurately describe the directional d- and f-electron bonding in these materials. Below is a comparison of snapshots of the final

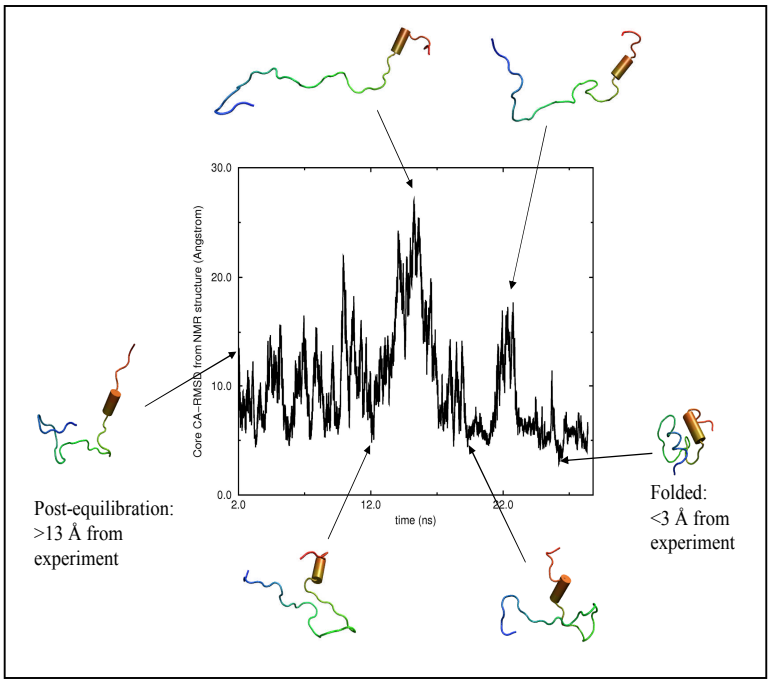
configuration in three different-sized simulations in which molten tantalum at 5000 K is isothermally compressed to a final pressure of 250 GPa, resulting in solidification to a polycrystalline phase. The 16-million-atom simulation on the right yields the first realistic grain-size distribution emerging entirely from the melt, reflecting our ability to model natural nucleation and growth processes as well as to create grain structures that reflect the directional quantum-mechanical bonding of the atoms—a result that was unobtainable without BlueGene/L and the early work done on Thunder.
— F.H. Streitz, M.V. Patel, J.N. Glosli



64k atoms (2048 processors) | 256k atoms (2048 processors) | 2M atoms (16384 processors) | 16M atoms (32768 processors)

1.3.3 Large, Long-Timescale Simulations of Protein-Folding Thermodynamics

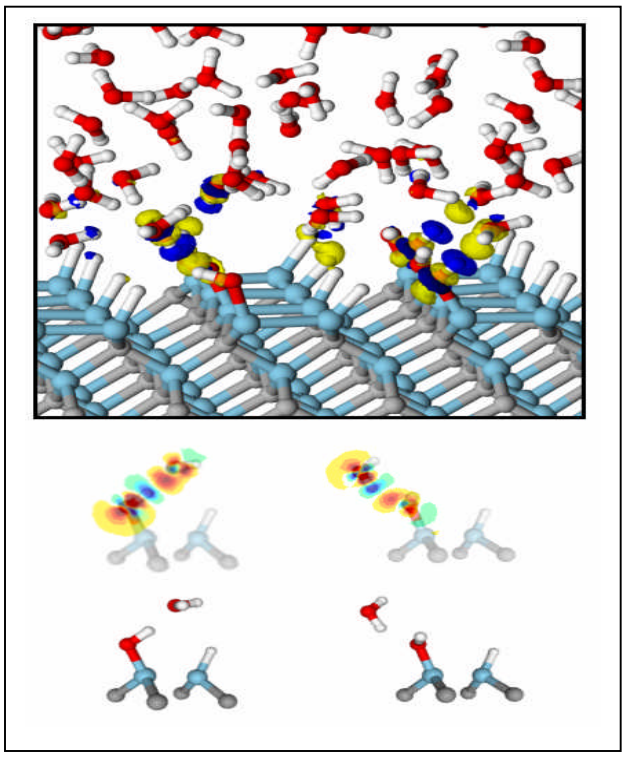
Thunder has enabled us to carry out the largest, longest timescale simulations of protein folding thermodynamics to date. Our simulations are helping to resolve an experimental controversy over the folding mechanism of the 40-amino-acid protein 1BBL—whether or not it is a barrierless "downhill folder." These simulations are roughly three times larger (47,000 atoms vs. 16,000 atoms) and run seven times longer (7 ms vs. 1 ms) than similar studies. Large system sizes are necessary to accommodate extended, unfolded structures, while long simulation times are necessary to sample from the folded state to the unfolded state and back. The figure shows one such transition from our Thunder runs, one of the first examples of reversible folding in a protein this large.



— Farid Abraham, Jed Pitera, Bill Swope

1.3.4 First-Principles Molecular Dynamics Simulations of Water

We have carried out a series of large-scale first-principles molecular dynamics simulations of water confined between hydrophobic and hydrophilic surfaces. Our preliminary results indicate that the precise nature of the confining media can induce large differences in the spatial distribution of water. Surprisingly, we find no evidence for the formation of ice-like layers near the surface and an overall tendency of hydrophobic surfaces to cause long-range ordering in the liquid. We have begun a separate investigation of the performance of density functional theory in reproducing the x-ray absorption spectrum of water and ice.
— Eric Schwegler



2. Capability and Capacity Requirements

This section provides a summary of the total capability and capacity requirements gathered to substantiate the requests made in this White Paper. Appendix 2 gives a complete list of actual requirements as they were collected from the scientists. M&IC provided the scientists with a template showing requirements collected from the survey in 2001 and asked them to respond by indicating any new or updated requirements. For the purposes of this survey, capability was considered to be calculations that would require access to the proposed 40 TF Peloton.[4] Capacity requirements may be thought of as representing day-to-day needs for runs in much greater quantity but at much smaller scale. This would include the small-scale runs necessary to develop the intuition required to prepare for meaningful capability runs.

First, to identify the intensive computing constituency, Table 1 provides a list of participants who responded to the White Paper request for information.

Table 1. May 2005 institutional survey participants.

	Project ID	Submitted By	Title	FTEs	Supports Lab Programs	Team From
1	BioImaging	Richard London	Complete Simulation of Single Molecule Imaging with an X-ray Free Electron Laser	2	LDRD, BIO	PAT/I, DNT/AX, BIO/MED
2	Z3	Barbara Lasinski, Bruce Langdon, Bert Still	Z3 Project	1	Area 2 of LLNL Long-Range Science & Technology Strategy; LDRD; A Prog; NIF	DNT/AX
3	BOUT/UEDGE	Xueqiao Xu, Tom Rognlien	BOUT (Boundary-Plasma Turbulence) Project	2	FEP	PAT/FEP
4	QCD/PHENIX	Ron Soltz	QCD/PHENIX	11	LDRD, other	PAT/N
5	pF3d	Bert Still, Dick Berger, Laurent Divol, Milo Dorr, Jeff Hittinger, Bruce Langdon, Steve Langer, Ed Williams	pF3d - Predictive Laser-Plasma Interaction Modeling	5.5	NIF, ASCI	DNT/AX, PAT/M, CASC
6	HP-CFD	Bob Lee, Don Ermak, Stevens Chan, Branko Kosović	High Performance Computational Fluid Dynamics (CFD) Models	3	EED, NAI, LDRD	EED/ASD
7	Climate	Philip Duffy	Climate Modeling	2	EED, other	EED/ASD
8	AtmosChem	Cyndi Atherton, Peter Connell, Cathy Chuang, Dan Bergmann, Philip Cameron-Smith	Global Atmospheric Chemistry	5	EED, other	EED/ASD
9	Homeland Security (Structures)	Charles Noble	Homeland Security (Structures)	3	HSO/HOPS	Eng/NTED
10	GEODYN/ RAPTOR	Benjamin Liu	AMR Shock Hydrocode with Coupled Strength and Radiation	12	A Program, Containment, Weapons Effects	EED, DNT/A

[4] Capability requirements are grand challenge-type simulations that may require a significant amount of time across half or more of the system with multiple ancillary runs requiring at least one-quarter of the system.

	Project ID	Submitted By	Title	FTEs	Supports Lab Programs	Team From
11	LDEC	Joseph Morris	Response of Underground Structures to Dynamics Loading	2.5	EED, LDRD, DNT/Q	EED
12	GEODYN/L	Tarabay Antoun	Unstructured Lagrangian Shock Hydrocode with Strength	1	Containment, Weapons Effects	EED
13	HighPlowZ	Eric Schwegler	Materials under Extreme Conditions	4	LDRD, PDRP, ASC	PAT/H
14	Compnano	Giulia Galli	Computational Nanoscience	4.5	LDRD	PAT/H, CMS
15	Beryllium	Eric Schwegler	Ab Initio Determination of the Beryllium Melting Curve	2	NIF/ICF	PAT/H
16	Nanosim	Giulia Galli	Nanoscale Fluid Flow and the Properties of Confined Water	3	LDRD	PAT/H
17	EvRec	Branko Kosovic	Dynamic Data-Driven Event Reconstruction for Atmospheric Releases	5	EED, LDRD, other	EED/ASD, Eng/DSED, Comp/CASC/CAR
18	Tempest	Milo Dorr	Kinetic Simulation of Boundary Plasma Turbulent Transport	6.6	FEP	PAT, Comp, CMS
19	Folding	Farid Abraham, Olgica Bakajin, Jed Pitera and William Swope	Discovering the Folding Rules that Proteins Obey	1?	LDRD, GTL	CMS
20	MDHydro	Alison Kubota and Wilhelm G. Wolfer	Hydrodynamic Instabilities from Fully Atomistic Paradigms	1.5	ASC, NIF, B Program	CMS
21	Rad Effects	Alison Kubota, Wilhelm G. Wolfer	Radiation Effects in Stockpile Materials	0.75	ASC, B Program	CMS
22	Aging Effects	Alison Kubota, Wilhelm G. Wolfer	Dynamic Properties of Aging in Metals	1.5	ASC, B Program	CMS
23	VASP	Jess Sturgeon, Wilhelm G. Wolfer	Electronic Structure Calculations of Point Defect Properties and Alloying Agents in Materials	1.5	ASC, B Program	CMS, PAT
24	mesochem	Richard Gee, Larry Fried, Amitesh Maiti	Mesoscale Material Dynamics and Code Development for HE/PBX Materials	3	LDRD, ASC, B Program, other	CMS
25	ParaDiS	V. Bulatov	Plastic Strength from Dislocation Dynamics	3	ASC P&EM	CMS
26	QuantChem	Krishnan Balasubramanian	Quantum Chemical Modeling of Materals, Molecules, and Clusters	10	DNT-ESC, LDRD	CMS/CCHED
27	PIMCQVIB	Kurt R. Glaesemann	Quantum Vibrations in Molecules	1	CMS	CMS
28	ShockRad	Evan Reed	Characterization of Radiation Emission from Shocked Materials	1		URP, CMS
29	Poly	Naida Lačević, Richard Gee, Larry Fried	Polymer Crystallization, Data Analysis	1.5	LDRD, B Program, other	CMS
30	Computational Chemistry and Biology	Chris Mundy, Will Kuo, Nir Goldman	First-Principles Molecular Dynamics for Understanding of Rare Chemical Events in Chemistry, Materials Science, and Biology	3	EMC, DNT, NAI	CMS/CCHED

	Project ID	Submitted By	Title	FTEs	Supports Lab Programs	Team From
31	Superionic	Nir Goldman	First-Principles Simulations of Superionic Solids	1.5	LDRD	CMS
32	ShockMetal	Eduardo Bringa	High-Strain Rate Loading of Metals	0.8	NIF, Eng, DNT, PAT	CMS
33	AstroMat	Eduardo Bringa	Materials Sciences for Astrophysics	0.2	PAT (IGPP)	CMS, PAT(IGPP)

Table 2 and Figure 1 document (in tabular and graphical formats, respectively) total capability and capacity needs in TF per year, by directorate. These requirement totals were obtained by performing a reasonable conversion of each project request to total TF per year and then a following summation based on the home organization of the responding principal investigator. From this summation, one can conclude that LLNL has a total need for capability cycles that is almost four times what could be provided by a 40 TF capability system in one year and a total need for capacity cycles that is more than twice what could be provided by 40 TF of capacity systems in one year.

Table 2. M&IC capability and capacity needs per directorate.

Directorate	Capability Requirements (TF/yr)	Capacity Requirements (TF/yr)
DNT	36.7	11.5
PAT	23.3	9.6
EED	23.3	28.8
CMS	70.0	38.3
Totals	**153.3**	**88.2**

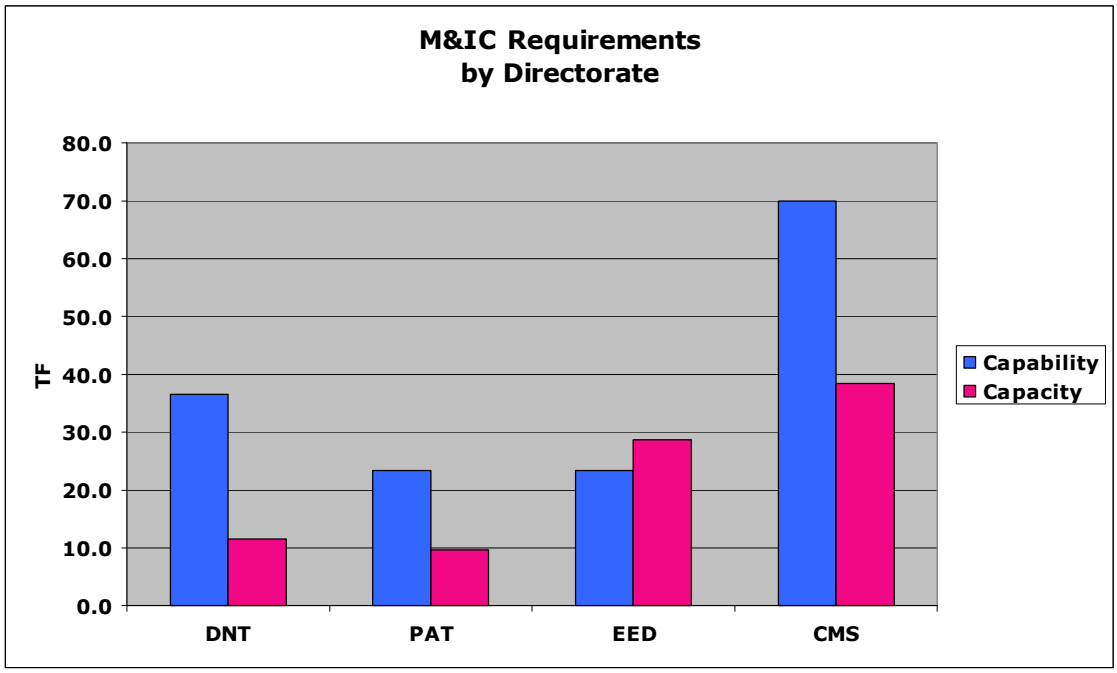


Figure 1. M&IC capability and capacity needs per directorate.

This presents the usual conundrum associated with infinite requirements and finite budgets. Experience has shown that this is most easily addressed by an approach that: (1) carefully allocates time based on institutional priorities and quality of perceived work, and (2) provides those who do not prevail in the competition with an allocation that allows the effort to continue to improve its science and methods so that it has hope to eventually earn a major allocation.

Through this approach, requirements that exceed resources by a factor of two to four can be managed. If the disparity gets much larger than this, however, the constituency denied access approaches the totality of LLNL science effort.

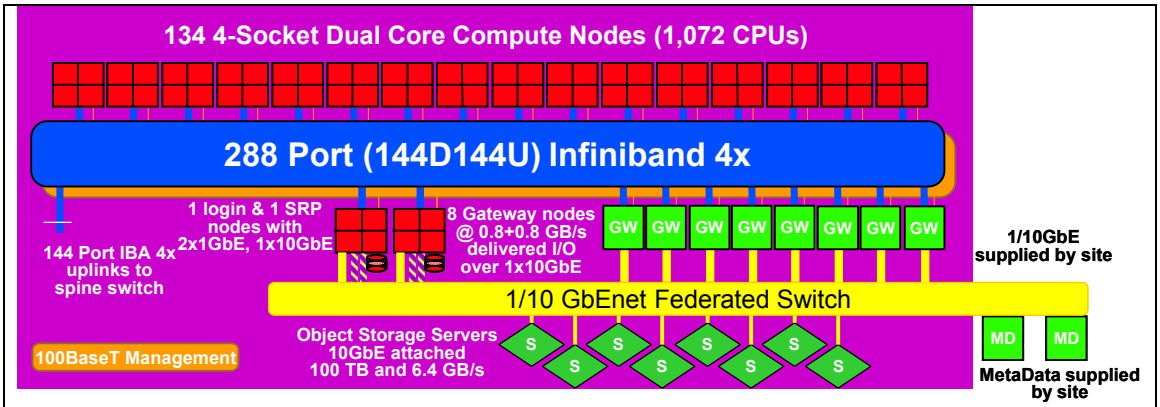
In summary, we are today at the point where the disparity between requirements and resources exceeds ability to manage the enterprise to good effect. With the resources and the vetting process proposed, the new systems we are requesting would be fully utilized and would position LLNL to continue the outstanding science and technology breakthroughs the M&IC program has enabled to date.

3. Recommended Architecture

The system will be built from scalable units (SUs), which consist of nearly complete computers in themselves, with the first level of the switch included in the unit. Clustering SUs, which are connected through the additional switch levels, creates larger computers. In this manner, a balanced system is fielded through a federated switch. The clustered system writes to the extensible, global parallel file system, Lustre.

Details of a particular design for an SU are shown in Figure 2. This example features AMD processors. The procurement process may permit a broader set of candidate solutions, so the precise processor and design of the SU will not be finalized until the procurement contract negotiation is complete.

One of the systems recommended in this proposal is for a single SU unit for smaller capacity jobs.



System Parameters

- 35.2 GF/s quad socket 2.2 GHz dual-core AMD (95W) SMP nodes with 16.0 GB, 25.6 GB/s DDR400 SDRAM (memory B:F=0.45, BW B:F=0.73).
- <3 µs, 2 GB/s MPI latency and bandwidth over IBA 4x (B:F=0.06).
- Support 800 MB/s transfers to archive over Jumbo Frame 10 Gb-Ethernet and IBA links from login node.
- No local disk. Local flash disk or remote boot and SRP target for root and swap partitions on RAID5 device for improved RAS; two login/service/master nodes.
- 400 MB/s POSIX serial I/O to any file system
- Disk capacity 100 TB (B:F=20) global parallel file system in multiple RAID5.
- I/O bandwidth 6.4 GB/s (B:F=0.0013) delivered parallel I/O performance.
- Software for build and acceptance: three provided software stack distributions (all based on several common key components: Red Hat, Lustre and Panasas, OpenIB, MPICH2, or OpenMPI).

Figure 2. Example SU system architecture for 144 nodes, 5.1 TF/s peak.

As stated previously, SUs can be aggregated. The design shown in Figure 3 represents a collection of eight SUs, a reasonable approximation for the capability system recommended in this document.

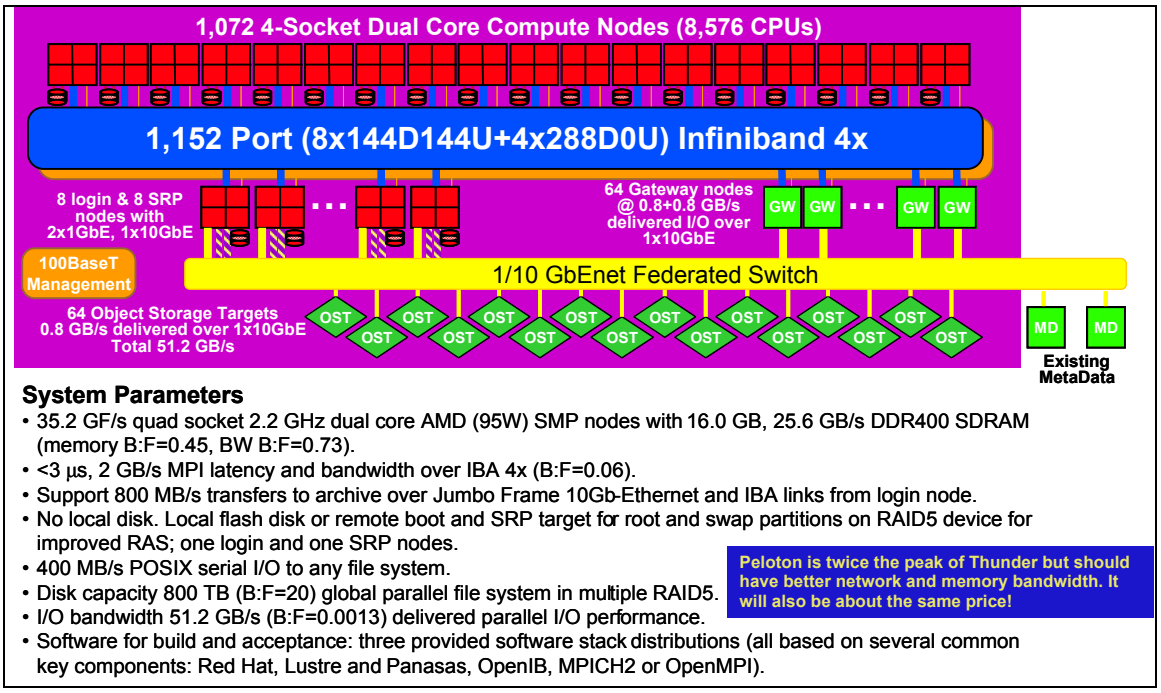


Figure 3. Peloton system architecture for 1,152 nodes, 40.6 TF/s peak.

4. Costs and Timeline

M&IC proposes to procure an additional 45 TF for the institutional computing community. Our strategy will be based on new 8-way SMP nodes that could be clustered into 5 TF scalable units (SUs). The cost estimate for each ~144-node SU is $2 million. This $2 million consists of $1.75 million for Peloton vendor costs (nodes, disks, and interconnect) and $250K per SU for costs to site systems, software licensing, and to maintain a balanced infrastructure (visualization, archive, networks). There is an additional cost of $1.5 million to cover the Peloton procurement and lease-to-own (LTO) costs.

For the computers, in total, we are requesting $19.5 million (9 SUs × $2 million + $1.5 million) over FY06–08 to procure the 45 TF Peloton (option no. 1 in Table 3). An additional $1.1 million in yearly ongoing base funding will be requested to pay for increased fixed costs associated with the new system. Eight of the SUs will be combined to form our new capability system, and the remaining SU would join Thunder, MCR, and GPS/iLX as our capacity systems.

This procurement is scalable; a smaller system could be procured at a proportionately lower cost (see Table 3). For instance, if the peak of the capability system were to be reduced by 10 TF (two SUs) to 35 TF, the cost would be reduced (by ~ $4 million) to $15.17 million. These numbers represent estimates; final costs will not be known until the procurement contract is complete.

We are also procuring a disk farm (collaborative data oasis) to be placed on the "green" network. We have received $850K in FY05 IGPE Capital Funds to procure the collaborative data oasis hardware.

Table 3. Cost options for procurement of Peloton computer.

Option	TF/s	Scalable Units	Computer Cost ($M)	Financing	Computer Out-years Total ($M)
1	45	9	18	1.50	19.50
2	40	8	16	1.33	17.33
3	35	7	14	1.17	15.17
4	30	6	12	1.00	13.00
5	25	5	10	0.83	10.83

The annual request of the institution to cover the costs for a 45 TF Peloton cluster is shown in Table 4 and Figure 4 (in tabular and graphical formats, respectively).

Table 4. Annual costs for option no. 1 from Table 3 and related expenditures.

Costs ($K)	FY05	FY06	FY07	FY08
Base - Fixed	5,360	6,460	6,460	6,460
Incremental - One Timers	916	7,058	6,500	6,500
IGPE	1,195	–	–	–
Program Contributions	973	1,118	885	885
Total Costs ($K)	**8,444**	**14,636**	**13,845**	**13,845**

Note: FY06 increase in Base is to pay ongoing fixed costs of the Peloton cluster.

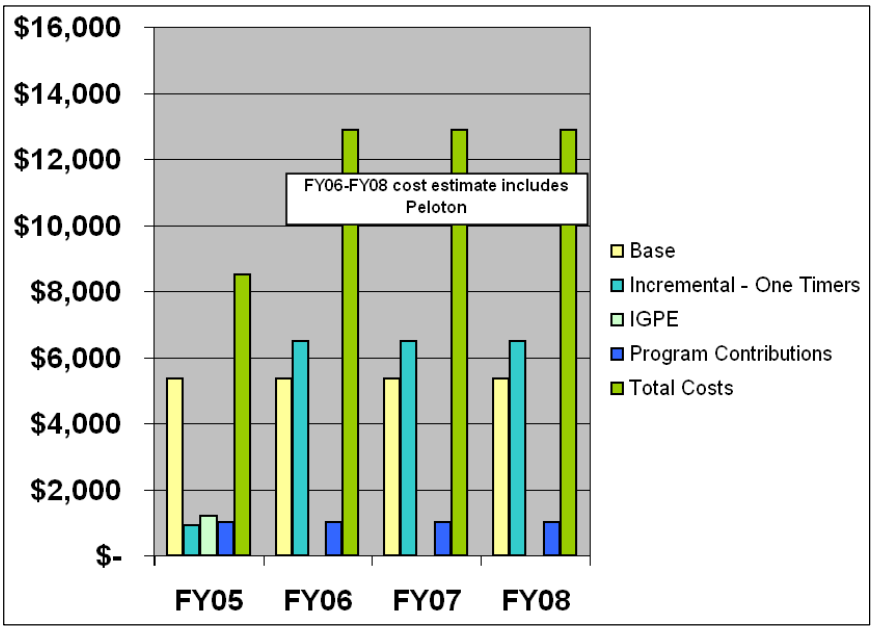


Figure 4. Annual costs (in $K) for option no. 1 from Table 3 and related expenditures.

Our desired timeline for the procurement and use of the 45 TF system is shown in Figure 5. The required milestone dates to meet both the desired science timeline and the Peloton timeline are included for CY05 through Q1-CY07.

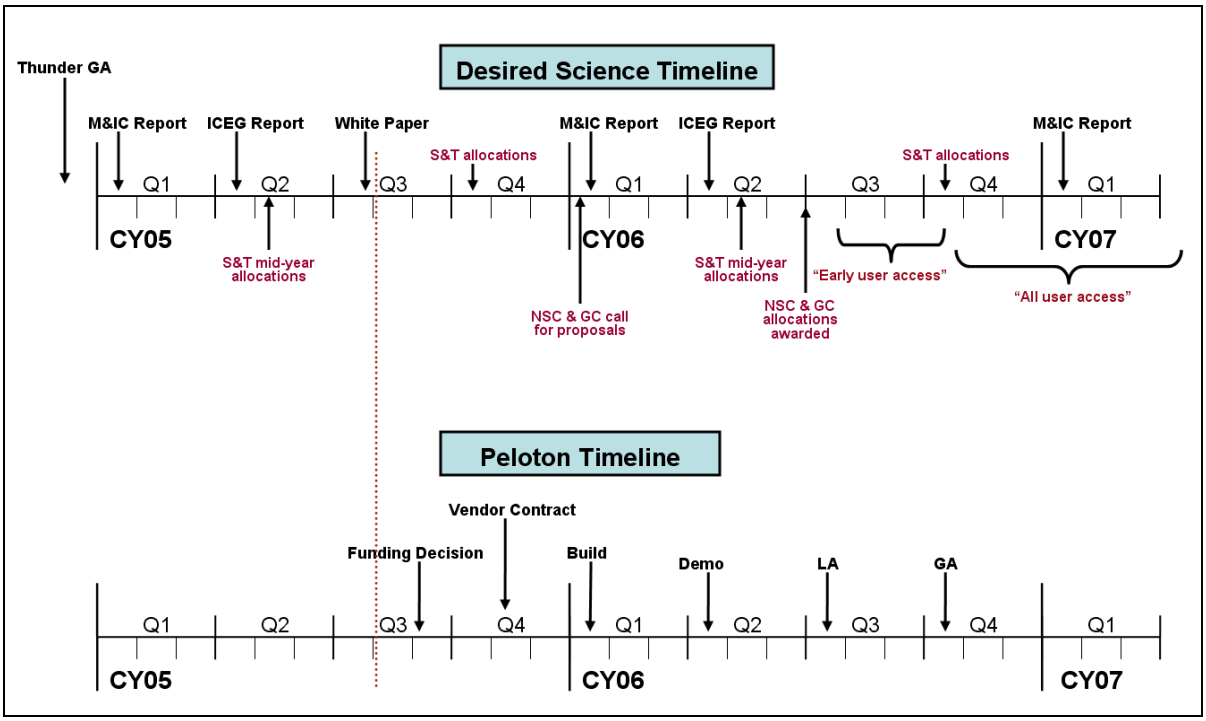


Figure 5. Timeline for procurement and use of the 45 TF system.

5. History of M&IC, Governance Model, and Strategy

The discussion that follows provides background for those interested in the history of Multiprogrammatic and Institutional Computing (M&IC) and its drivers. It also provides additional motivation to better understand M&IC strategies and current challenges. In the latter half of the 1990s, the maturation of parallel computing technology made it possible for the nation to contemplate the development of production-level 3D scientific applications requiring super-teraflop computational capability. In fact, the Stockpile Stewardship Program led in identifying this potential and proposed ASCI as its spearhead to enable certification, in conjunction with subcritical and other experiments and theory, in the absence of underground testing.

LLNL, as an institution, recognized that if one of its major programs was embarking on an adventure that had the potential to revolutionize how science would be done in the next century, the health of the institution depended on an S&T base that also had access to powerful ASCI-class computing environments. This was a strategic move that has kept the disciplines at the forefront and positioned LLNL as the preeminent simulation site today.

From this notion was born Multiprogrammatic and Institutional Computing. M&IC is truly institutional. Many directorates invest, and the institution invests. The growth of M&IC since 1997 has been significant (see Table 5 and Figure 6); the total capacity currently available to M&IC scientists is about 35 TF/s.

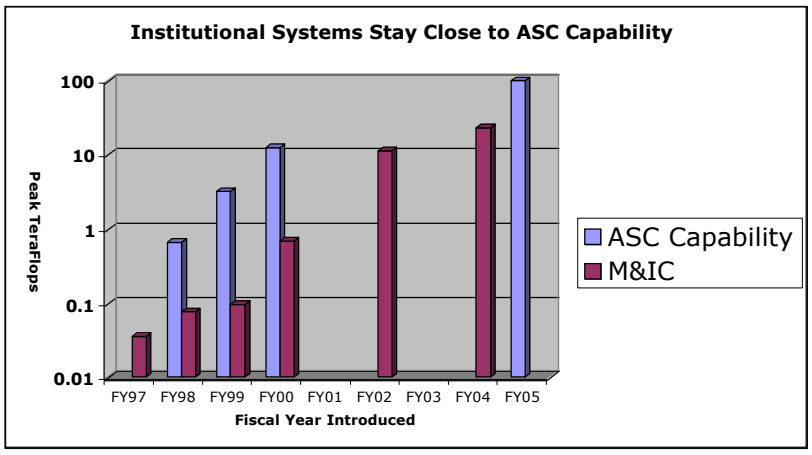


Figure 6. Comparison of M&IC to ASC capability from 1997–2005.

Table 5. Growth of M&IC computing power (in MF/s) from 1997–2005.

	FY97	FY98	FY99	FY00	FY01	FY02	FY03	FY04	FY05
T3D	37	0	0	0	0	0	0	0	0
Compass	35	70	70	70	70	0	0	0	0
TC98	0	0	96	108	176	147	0	0	0
SUN	0	12	24	12	12	12	0	0	0
Qbert	0	0	0	0	0	0	12	25	25
TC2K	0	0	0	683	683	683	683	683	0
GPS	0	0	0	0	0	192	277	277	277
LX	0	0	0	0	43	101	0	0	0
iLX	0	0	0	0	0	0	634	678	678
ASC Blue	0	16	89	99	74	74	0	0	0
ASC Frost	0	0	0	0	326	326	0	0	0
MCR	0	0	0	0	0	11059	11059	11059	11059
Thunder	0	0	0	0	0	0	0	22938	22938
Total	72	98	279	972	1384	12594	12665	35660	34977

M&IC is largely supported by institutional investments; however, about $1 million comes from investing programs across the Laboratory or directly from directorates. (Table 6 and Figure 7 show the recent M&IC cost history in tabular and graphical format, respectively, and Figure 8 shows the institutional ownership in M&IC resources.) The governance model is both grass roots and hierarchical. The "board of directors" (the Institutional Computing Executive Group, or ICEG) consists of well-known LLNL scientists who are qualified to point out deficiencies and request improvements. Typically, ICEG members are appointed by ADs in the various directorates. Hierarchically, M&IC management reports to the Director's Office, namely to the Deputy Director for Science and Technology, who provides guidance relative to the institution's overall S&T goals and at the highest level manages allocations. Generally, it is not difficult to meet both the scientists' requests and the institution's, and this is a challenge that M&IC facilitates. Lest the investment levels highlighted here be viewed as excessive, we note that the M&IC environment is comparable to the best unclassified environments anywhere in the country, and the total investment here is 25 to 33% that of NERSC or Oak Ridge. Such is the power of leverage and momentum coming from partnership with the ASC Program.

Table 6. M&IC costs for FY00-FY04, and FY05 funding.

Costs ($K)	FY00	FY01	FY02	FY03	FY04	FY05
Base - Fixed	2,044	2,055	3,055	4,383	4,208	5,360
Incremental – One Timers	1,684	3,417	3,300	7,118	9,062	916
IGPE	–	–	–	–	–	1,195
Program Contributions	935	935	935	935	1,868	973
Total Costs	**4,663**	**6,407**	**7,290**	**12,436**	**15,138**	**8,444**

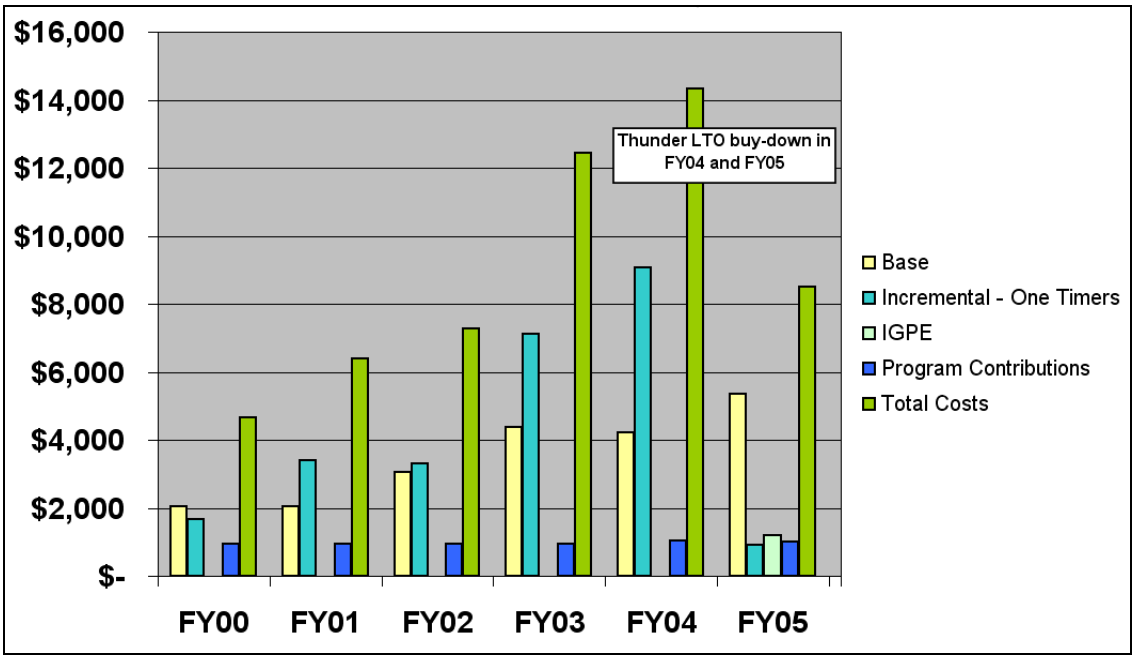


Figure 7. M&IC costs (in $K) for FY00–FY04, and FY05 funding.

The institution covers all the operational costs and also invests in the high-performance computing hardware. The programs and directorates invest only in the hardware. A share of the

computing resource—called a bank—is correlated to the level of investment. The size of the bank is proportional to the level of investment.

Access to the institution's banks is managed through a proposal process. This process is described at a high level in the Executive Summary. There, we recommend three modes of access for the future. Previously, there were only two modes of access. The additional mode will enable external user access to institutional hardware, as long as the winning proposals are aligned with S&T strategic objectives and are in collaboration with LLNL scientists. In the past, M&IC management worked with the ICEG to do initial reviews of proposals, passed on recommendations to the Deputy Director for S&T, and then implemented the Deputy Director's decision. The ICEG members typically manage programmatic banks, reflecting the priorities set by their respective ADs.

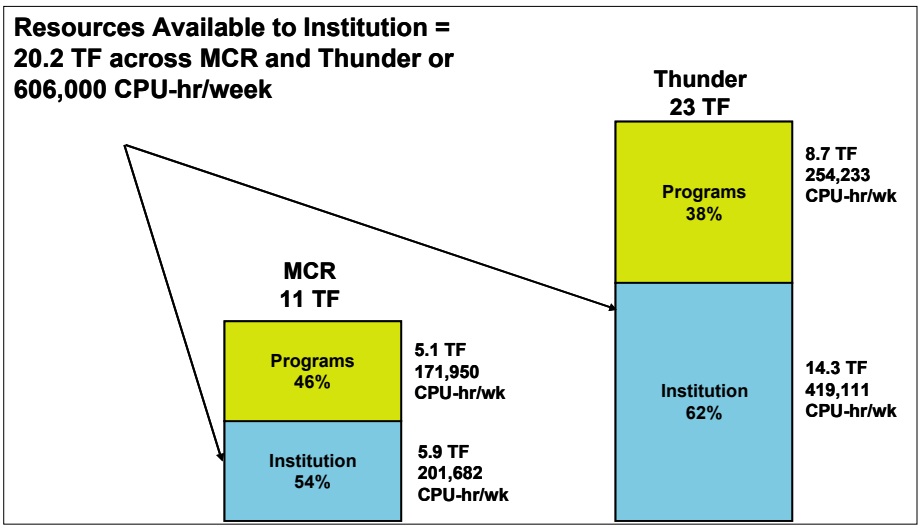


Figure 8. Institutional ownership in M&IC resources.

As was mentioned in the Executive Summary, the Director's Office elected to use Thunder for capability computing. The result of the proposal review process was that the winning projects were divided into two tiers (see Table 7), with Tier 1 receiving larger allocations and higher priority. A number of Tier 2 proposals were identified, and these gain access to the machine with smaller banks and slightly lower priority. By adding Tier 2 proposals, we hoped to avoid any idle periods on the machine. Our concerns of any significant idle cycles turned out to be unjustified.

Table 7. Thunder tier 1 and tier 2 projects.

Tier 1 Projects			
PI	**Organization**	**Bank**	**Project Description**
Bulatov, Vasily V.	CMS	lines	Dynamics and Self-Organized Patterning of Lines
Govindasamy, Bala	E&E	clchange	Regional Scale Climate Change Detection
Schwegler, Eric	PAT	nanosim	Wet Transport of Nanoparticles in Highly Confined Media
Mundy, Christopher J.	CMS	fph2o	Atmospheric and Planetary Water Using First-Principles Methods
Abraham, Farid Pitera, Jed Swope, William	CMS	folding	Protein Folding at the Speed Limit
Reed, Evan	CMS	shockrad	Emission of Coherent Radiation from Shock Waves

Tier 2 Projects			
PI	**Organization**	**Bank**	**Project Description**
Hood, Randolph Q. Patel, Mehul V.	PAT	squeeze	Hybridization and Multiphase Competition of High-Z Metals
Duffy, Philip B.	E&E	chmodel	Predictability of Precipitation in California during El Niño Winters [LDRD]
Galli, Guilia	PAT	compnano	Atomically Controlled Synthetic and Organic Nanostructures [SI]
Schwegler, Eric	PAT	chemd	Chemical Dynamics at Interfaces [LDRD]
Still, Bert	AX	z3d	Particle-in-Cell Modeling of Short-Pulse High-Intensity Laser Plasma Interactions

Thunder and MCR are today both saturated (see Figure 9). Two related factors are emerging that threaten M&IC's continued success. The first is that a number of LLNL teams required significantly larger levels of access than can be granted. A number of high-quality projects received zero capability allocations. A check of the queues on MCR (capacity) and Thunder (capability) at any particular time shows many jobs waiting, with some in this state for several days. The second factor is that M&IC capability systems will very soon be uncompetitive with the best unclassified systems outside of LLNL (for example, Sandia's new institutional system will be 59 TF/s). This will erode the advantages accruing to our scientists.

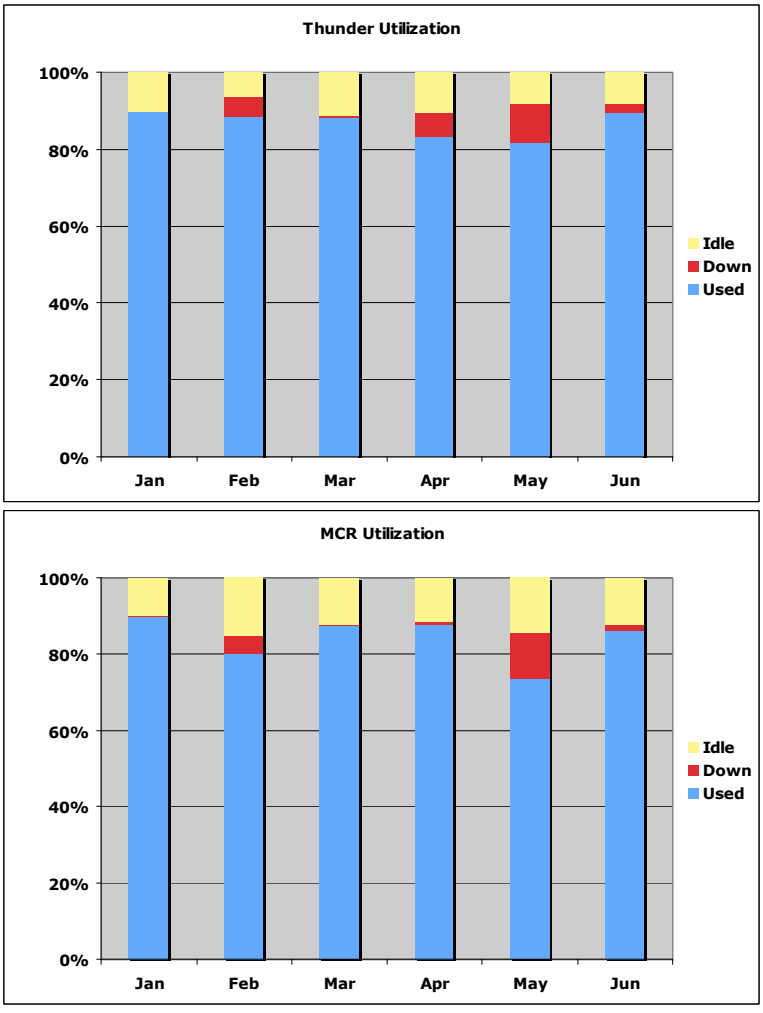


Figure 9. Utilization for the first six months of 2005 reveals saturation of capacity (MCR) and capability (Thunder) resources.

In responding to this challenge, LLNL is not without its strengths. ASC has provided the R&D funds that make it possible for M&IC to invest in high-performance but low-cost infrastructure systems, usually before most other sites can do so. Again because of ASC investments, LLNL has the benefit of one of the most experienced and well-staffed scientific computing centers in the world. An investment in hardware is leveraged by attention from experienced integrators, operators, and services staff, and from a well-engineered foundation in networks and storage. All of this mitigates considerably the risks inherent in investing in the newest and best cost performance technologies.

Our platform strategy has been to straddle multiple technology curves to appropriately balance risk and benefit. We follow three complementary technology curves, as shown in Figure 10. The first allows us to deliver to today's stockpile needs, the second delivers an affordable path to a future petaflop system, and the third provides a low-cost transition to the next generation of platform. M&IC investments have favored curve #3, open source commodity clusters, and in fact have led the way worldwide until recently. We believe that for the next 2–3-year cycle, clusters are the best solution for M&IC. In 2008, we will decide between clusters and an advanced technology solution (curve #2) for M&IC, depending on the maturity of technologies at that time.

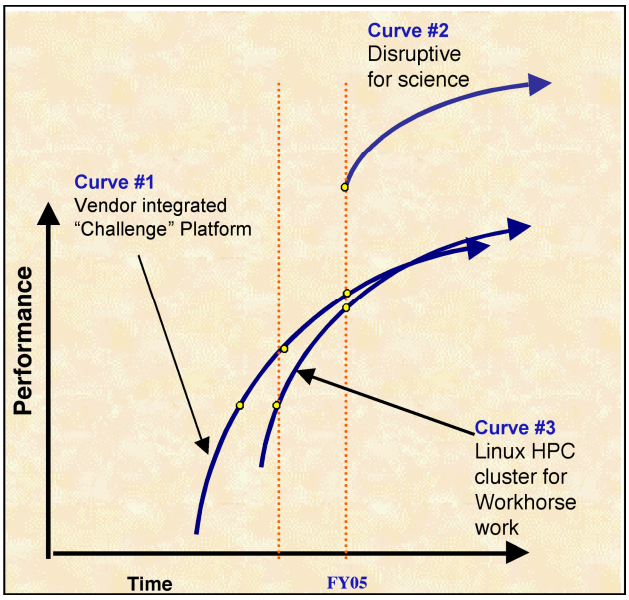


Figure 10. Multiple technologies foster viable paths to the future while delivering necessary computing cycles today.

6. Summary

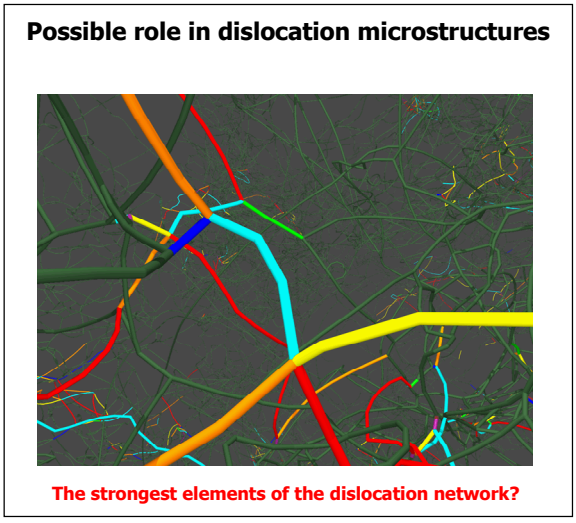
We have described a strategy and made a request that would maintain M&IC as a preeminent and balanced facility for science by **expanding capability while maintaining a 1:1 TF ratio of capacity to capability**. Further, we have advanced the notion of facilitating external collaborations through **National Security Challenge** proposals that offer capability computer access to external collaborators interested in working with LLNL scientists in strategic thrust areas. Last, we have described a **collaborative data oasis** to facilitate collaborations with scientists from DOE-designated sensitive countries.

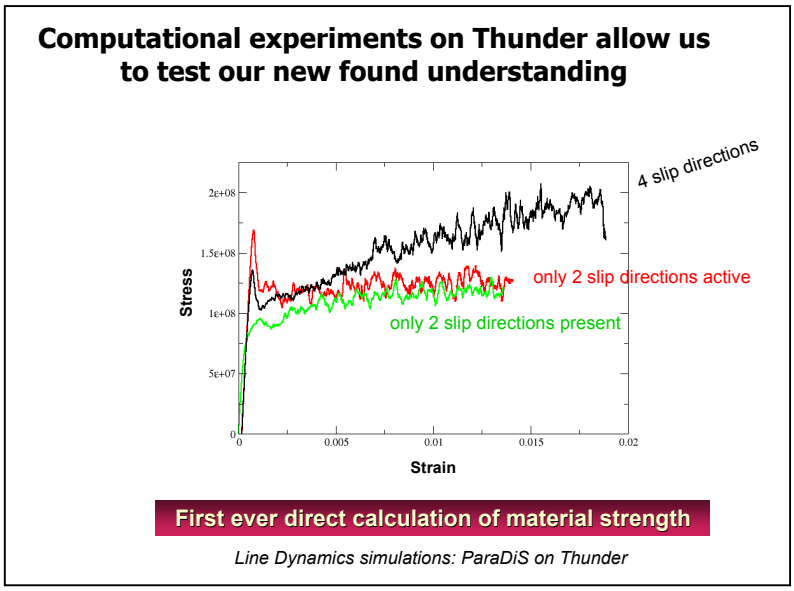
We have received $850K in FY05 IGPE Capital Funds to procure the disk farm necessary to create an approximately 1 PB collaborative data oasis. We are requesting an additional $19.5 million across FY06, FY07, and FY08 to procure a large, 40 TF/s capability system and a 5 TF capacity system for small jobs. Existing computers (Thunder and MCR) would be used to fill in the capacity gap between 40 TF and 5 TF. The newly procured systems would be financed under a lease-to-own (LTO) in FY06–08. The procurement of a new 45 TF Peloton cluster would also require an additional $1.1 million per year be added to the ongoing base funding for M&IC to pay for increased fixed costs associated with the new cluster.

Appendix 1. Impacts and Results as Described by Scientists

A1.1 Dislocation Dynamics in BCC Metals

We have developed a new understanding of the topology of dislocation networks in BCC metals. Each given line can be traced from one multinode to the next, but it loses itself in the multinode. In between two multinodes, the line can also zip a few binary junctions. The multinodes are very strong obstacles and, at the same time, generate new dislocation density and contribute to strain hardening. In BCC, a regular network, consisting of binary junctions only, can be uniquely deconstructed into its constituent lines. However, as soon a single multinode forms, this topological memory is lost: while the network can still be deconstructed, there is no unique way to do so. The number of possible variants increases combinatorially (exponentially) with the number of multinodes. This picture is one of many that we obtained by finding multinodes in simulations performed on Thunder.
— Vasily Bulatov



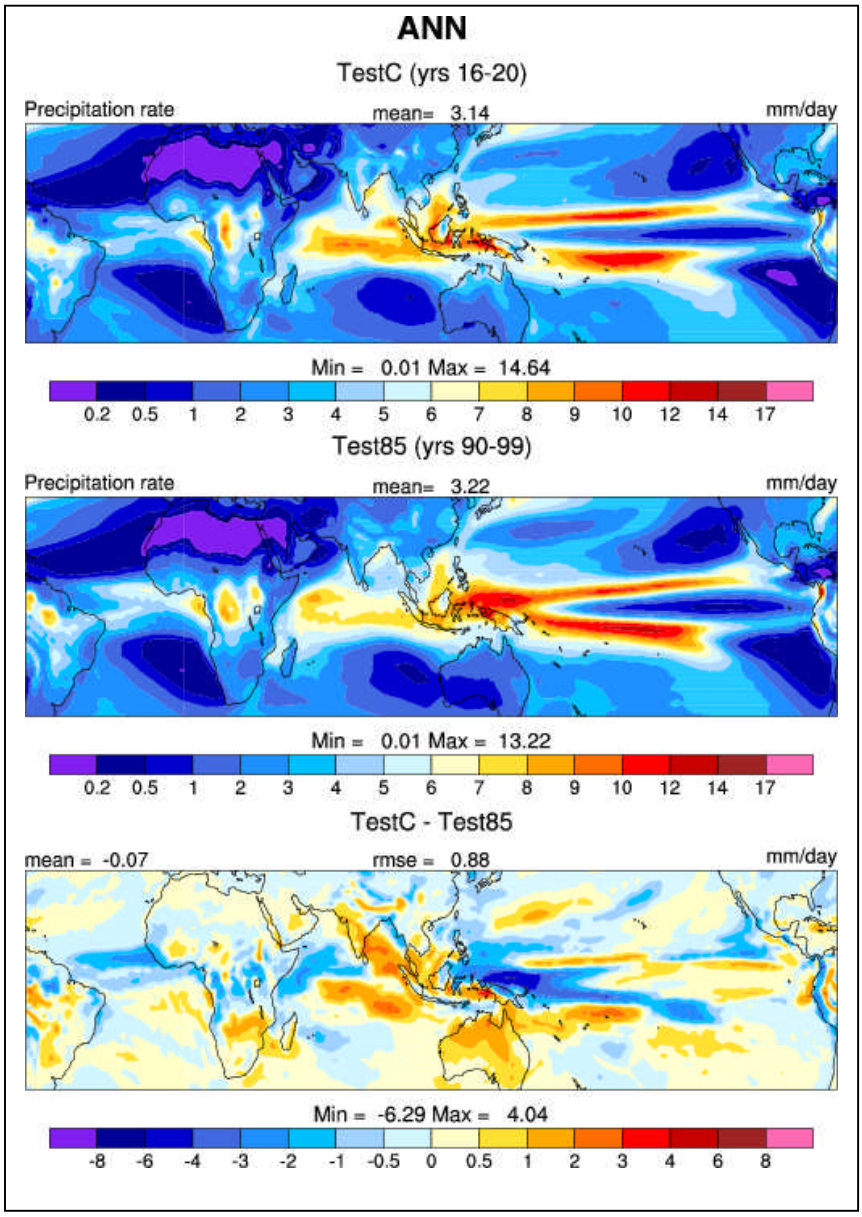


The black and red curves reproduce the experimental behavior observed in BCC metals where, when deformed along the direction of the highest symmetry (black), the crystal hardens at a high rate (the rate is equal to the slope of the curve). However, when the material deformed along a low symmetry direction, very little hardening is observed (the curve is nearly flat). The green line is a computer-aided thought experiment that cannot be done in a laboratory. Here we "rob" the system of two of its four possible slip directions, otherwise doing exactly the same simulation as black. The result is strikingly different. There is essentially no hardening; hence, because the only difference is the lack of multi-junctions (two remaining Burgers vectors do collide but can never form a multi-junction), we conclude this through experiment (done on Thunder) by stating that the high hardening rate observed in the black case is due to multi-junctions. This is a rather neat example that computational (and thought) experiments do not have to reflect reality to be useful for understanding.

A1.2 High-Resolution Climate Modeling on Thunder

We ported CAM3 (Community Atmospheric Model) and CCSM3 (Community Climate System Model) to Thunder. We performed 20-year CAM3 simulations on Thunder at 2×2.5 and 1×1.25 and 0.5×0.625 degree resolution using the Finite Volume (FV) dynamical core option. These results have been validated against observations and other control simulations performed at the National Center for Atmospheric Research (NCAR).

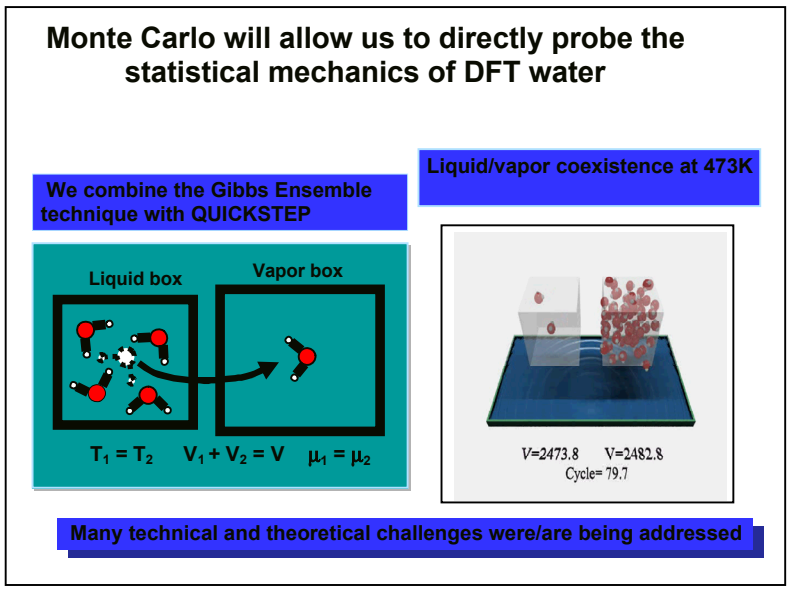
We also performed a 100-year control CCSM3 simulation at spectral T85 (1×4×1.4 degree resolution), and many 25-year CCSM3 simulations at the FV 2×2.5 degree resolution. It takes about 2 wall-clock hours on about 64 nodes to perform 1 year of model simulation. The FV 2×2.5 simulations were performed to understand the sea ice dynamics in the FV option. A 400-year CCSM3 simulation at the FV 1×1.25 degree resolution is currently in progress.
— Bala Govindasamy



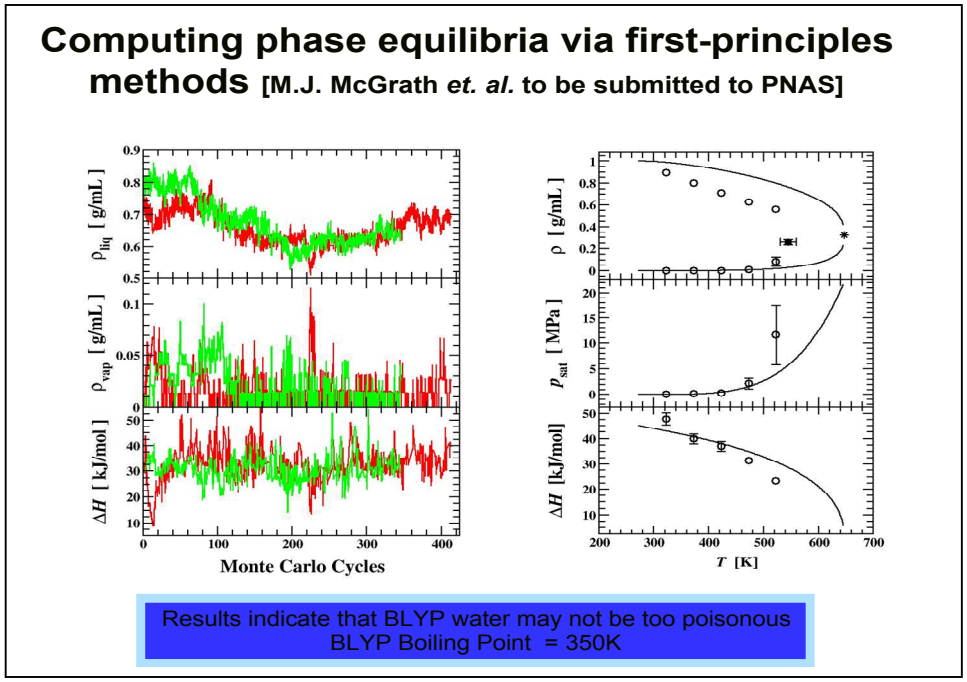
This plot compares annual mean tropical precipitation from the CCSM3 simulations performed at FV 1x1.25 degree resolution (TestC; top panel) and a spectral T85 (~1.4x1.4 degree resolution; Test85; middle panel). The bottom panel shows the difference. The precipitation pattern in the TestC simulation is similar to the control Test85. Both simulations were performed on Thunder.

A1.3 Monte Carlo Simulations of Phase Equilibria in Water

The movie snapshot from simulations made possible because of Thunder is the first two-phase calculation performed via first principles. You can see the liquid/vapor coexistence and from this determine all the thermodynamic properties. This has never been done directly before.
— Christopher Mundy

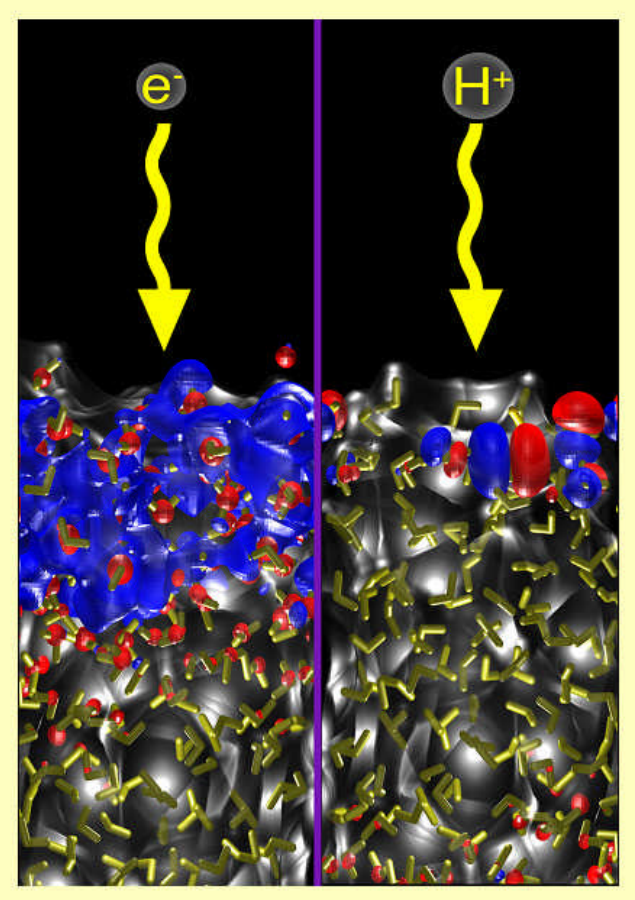


First-principles Gibbs ensemble Monte Carlo simulations for water based on density functional theory yield a qualitatively correct description of the vapor/liquid coexistence curve and reveal striking changes of the structural and electronic properties along the saturation line. Efficient Monte Carlo algorithms and a mixed-basis set electronic structure program were used to compute the vapor/liquid coexistence curve of water from first principles. A water representation based on the Becke-Lee-Yang-Parr exchange and correlation functionals yields a saturated liquid density of 900 kg/m^3 at 323 K, and normal boiling and critical temperatures of 350 and 550 K, respectively. An analysis of the structural and electronic properties of the saturated liquid phase shows an increase of the asymmetry of the local hydrogen-bonded structure despite the persistence of a four-fold coordination, and decreases of the molecular dipole moment and of the spread of the lowest unoccupied molecular orbital with increasing temperature.



A1.4 First-Principles Calculations of the Physical Properties of Water

This work used a "new" method (Monte Carlo with Electronic Structure calculations) to calculate, from first-principles, the properties of water. (It's "new" because no one has had the computing horsepower to do this.) The work in this area completed on MCR by Mundy and Will Kuo has been published. [See the bibliographic citation, item 6 in Section 1.2.]
— Christopher Mundy

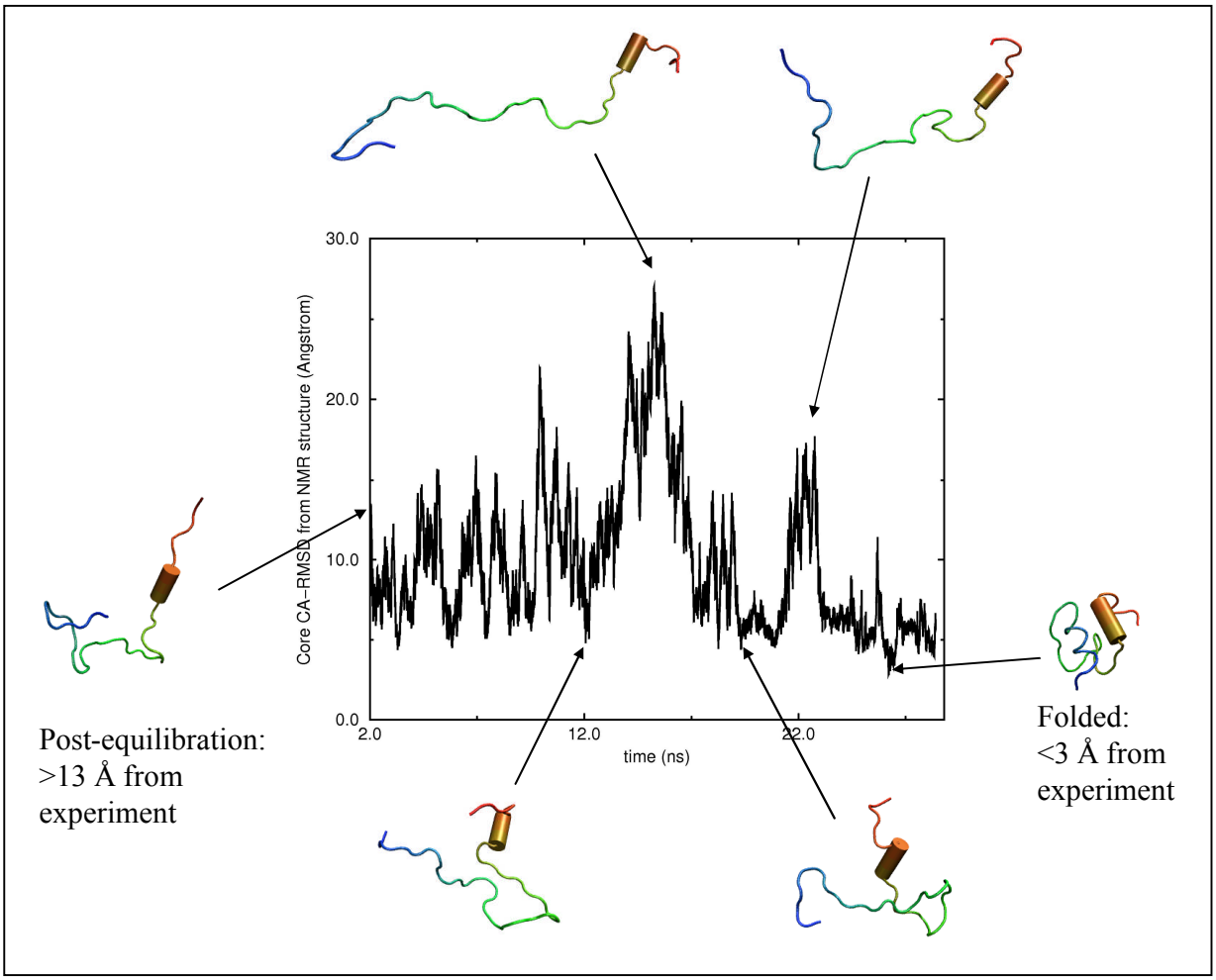


Water's properties are well known. If one can verify the method using water as a benchmark, then one can use the method to find the physical properties of nasty compounds that you really don't want to work with experimentally, such as organic phosphates (like sarin, the nerve gas). This work used up to 512 CPUs for 336 hours (equivalent to 84 hours on the full MCR machine). The figure shows the highest occupied molecular orbital on the surface as well as the lowest unoccupied molecular orbital. Both are on the surface, meaning the surface is reactive to both protons and electrons.

A1.5 Large, Long-Timescale Simulations of Protein-Folding Thermodynamics

Thunder has enabled us to carry out the largest, longest-timescale simulations of protein folding thermodynamics to date. Our simulations are helping to resolve an experimental controversy over the folding mechanism of the 40-amino-acid protein 1BBL—whether or not it is a barrierless "downhill folder." These simulations are roughly three times larger (47,000 atoms vs. 16,000 atoms) and run seven times longer (7 ms vs. 1 ms) than similar studies. Large system sizes are necessary to accommodate extended, unfolded structures, while long simulation times are necessary to sample from the folded state to the unfolded state and back. The figure below shows one such transition from our Thunder runs, one of the first examples of reversible folding in a protein this large.

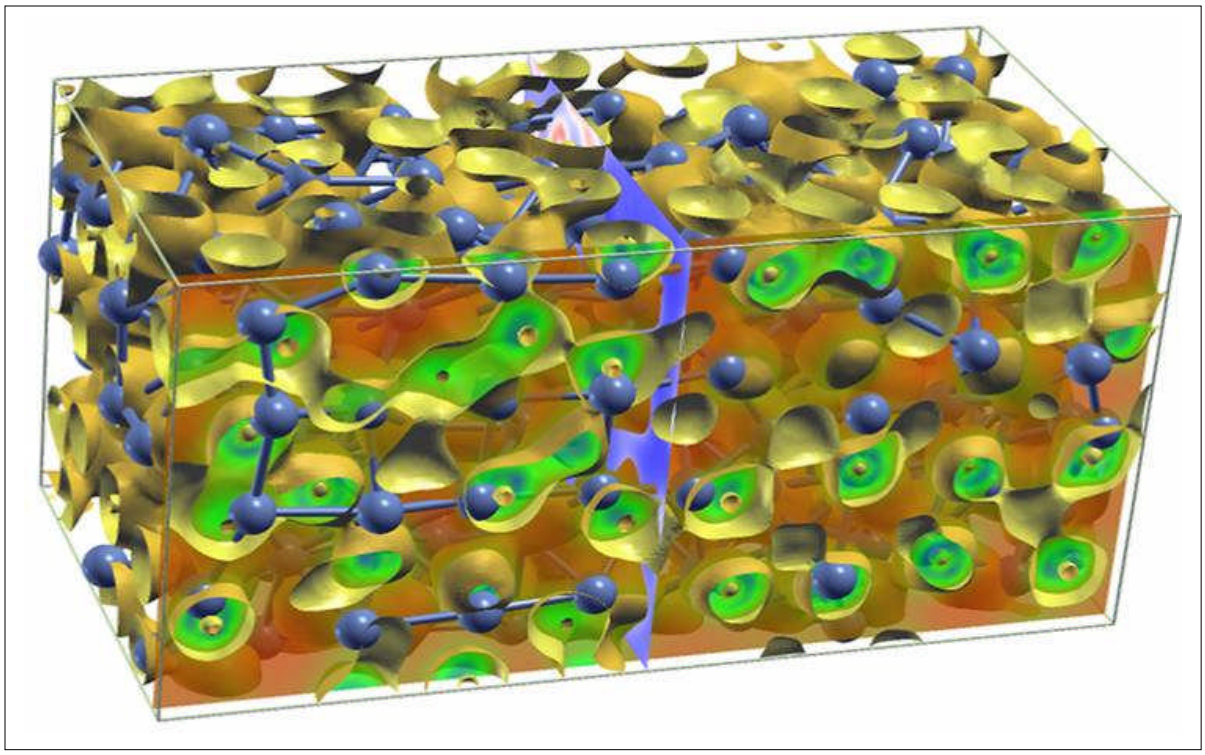
— Farid Abraham, Jed Pitera, Bill Swope



A1.6 Calculating the Melting Curve of Molybdenum

Eight of our planned 15 runs are under way to calculate the melting curve of molybdenum. Of these eight runs, five are near zero pressure and three are near 1 Mbar. These eight runs are 14%, 16%, 16%, 20%, 20%, 20%, 25%, and 35% completed.

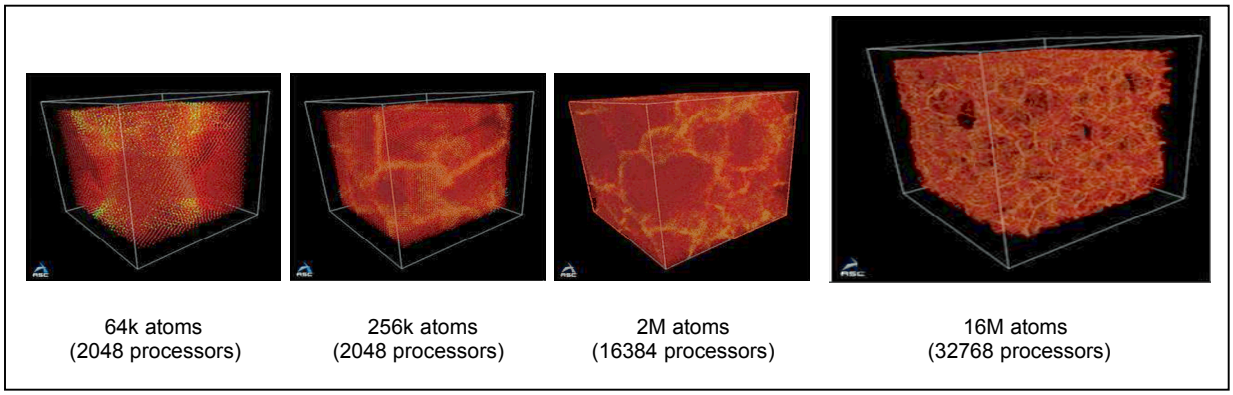
Below is a snapshot taken from one of these two-phase melting calculations. The positions of the atoms and nearest neighbor bonds are shown in blue. Electronic charge isosurfaces are also shown. The atoms on the left were initially in the BCC solid phase, while the atoms on the right were initially in the liquid phase.
— Randolph Q. Hood



A1.7 Molecular Dynamics Simulations of Pressure-Induced Resolidification in Tantalum

The short-range atomic structure and long-range grain structure created in a metal through rapid resolidification at high pressure and temperature are unknown, as is the timescale for the liquid-to-solid transition. Understanding the kinetics of resolidification and the morphology of the resolidified solid is vital to the accurate multiscale modeling of the mechanical properties of materials subject to extreme dynamic loading. We are performing molecular dynamics (MD) simulations to model the rapid resolidification of complex transition and actinide metals using advanced quantum-based many-body MGPT potentials that accurately describe the directional d- and f-electron bonding in these materials. Prior to the availability of BG/L, the added computational expense of these potentials limited us to 1 ns MD/MGPT resolidification simulations of about 10,000 atoms (requiring a month of run time), which were still approximately two orders of magnitude smaller in size than what was needed to access the required physics. Simulations very recently completed on BlueGene/L with LLNL's new parallel ddcMD code conclusively demonstrate that this supercomputer will enable us to overcome these artificial size limitations.

Below is a comparison of snapshots of the final configuration in three different-sized simulations in which molten tantalum at 5000 K is isothermally compressed to a final pressure of 250 GPa, resulting in solidification to a polycrystalline phase. The 16-million-atom simulation on the right yields the first realistic grain-size distribution emerging entirely from the melt, reflecting our ability to model natural nucleation and growth processes as well as to create grain structures that reflect the directional quantum-mechanical bonding of the atoms—a result that was unobtainable without BlueGene/L and the early work done on Thunder.
— F.H. Streitz, M.V. Patel, J.N. Glosli



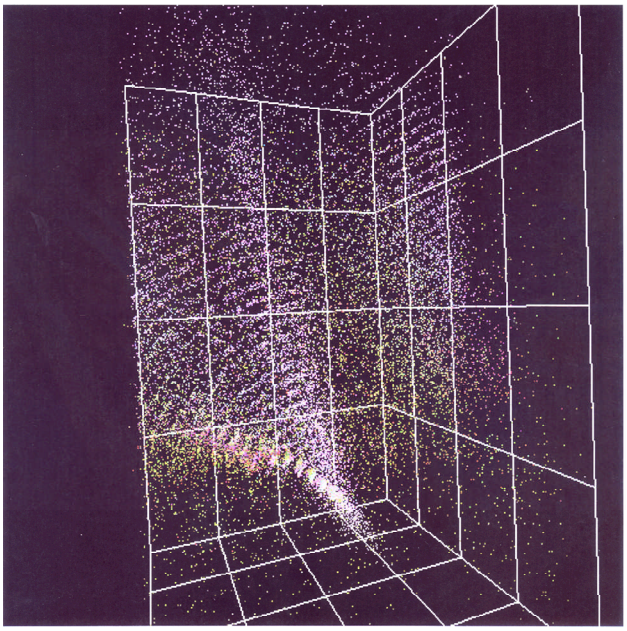
Four images of resolidified tantalum samples of increasing size after 0.75 ns of MD/MGPT simulation time. The 64,000-atom sample is seen to essentially consist of two large grains, with a grain boundary that spans the periodic simulation cell. This represents the creation of an infinite grain boundary—an unphysical result. The simulation performed with 256,000 atoms exhibits more realistic grain boundaries but with an artificially reduced variation in the resulting grain size distribution. The 2-million-atom simulation yields a more realistic distribution of grain sizes. The 16-million-atom simulation yields a much more realistic distribution of grain sizes.

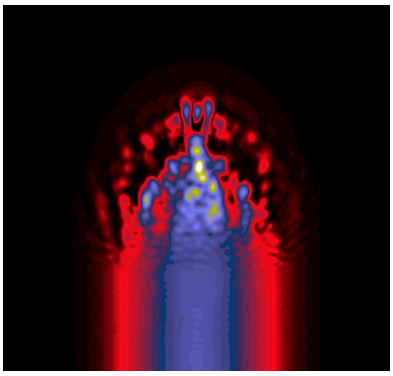
A1.8 Three-dimensional Visualizations of Laser–Plasma Interaction

We are exploring 3-D visualization of the structure of energetic electron flows. On Thunder, we have begun 3-D studies of laser–plasma interactions with "realistic" beam profiles.
— Barbara Lasinski, Bruce Langdon, Bert Still

oblique incidence; $16n_c$; 10^{19} W/cm^2

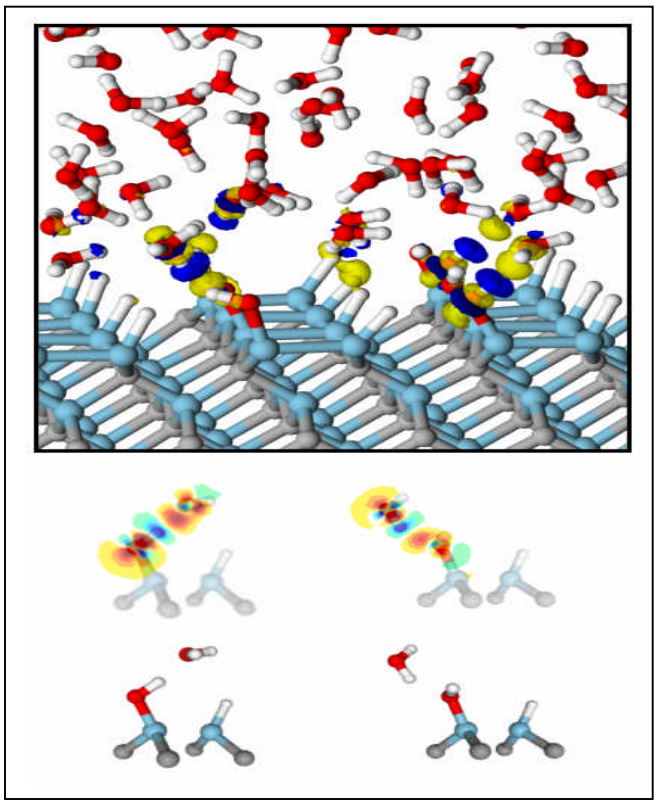
- oblique at 30° in x, and normal in y onto a $24\lambda_0 \times 24\lambda_0 \times 41\lambda_0$ plasma
- Jets of energetic electrons are seen in plots of (x, y, z) positions of electrons with energy > 4 MeV at 1 ps.
- The harmonic structure (striations) is a signature of the acceleration process.
- Particles are color coded by their directed velocities.
- This plot was made with yorgl (Steve Langer's GL extensions to yorick).





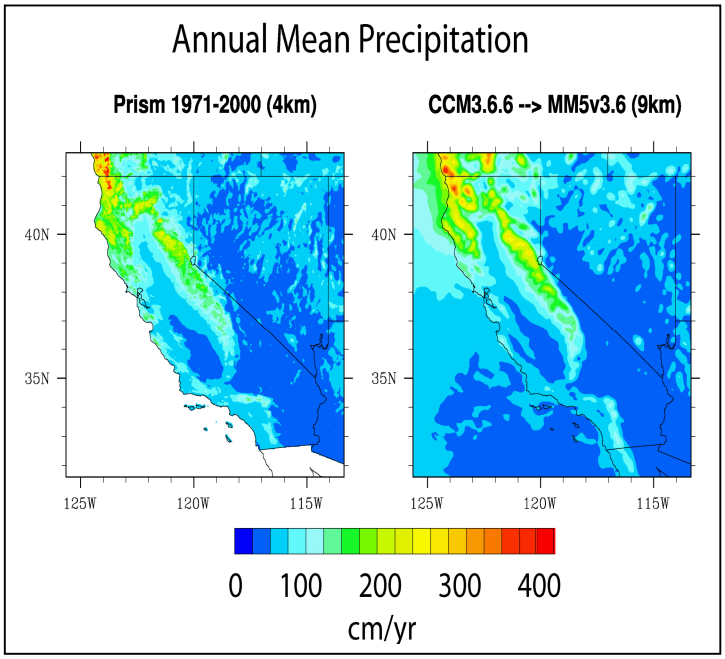
A1.9 First-Principles Molecular Dynamics Simulations of Water

We have carried out a series of large-scale first-principles molecular dynamics simulations of water confined between hydrophobic and hydrophilic surfaces. Our preliminary results indicate that the precise nature of the confining media can induce large differences in the spatial distribution of water. Surprisingly, we find no evidence for the formation of ice-like layers near the surface and an overall tendency of hydrophobic surfaces to cause long-range ordering in the liquid. We have begun a separate investigation of the performance of density functional theory in reproducing the x-ray absorption spectrum of water and ice.
— Eric Schwegler



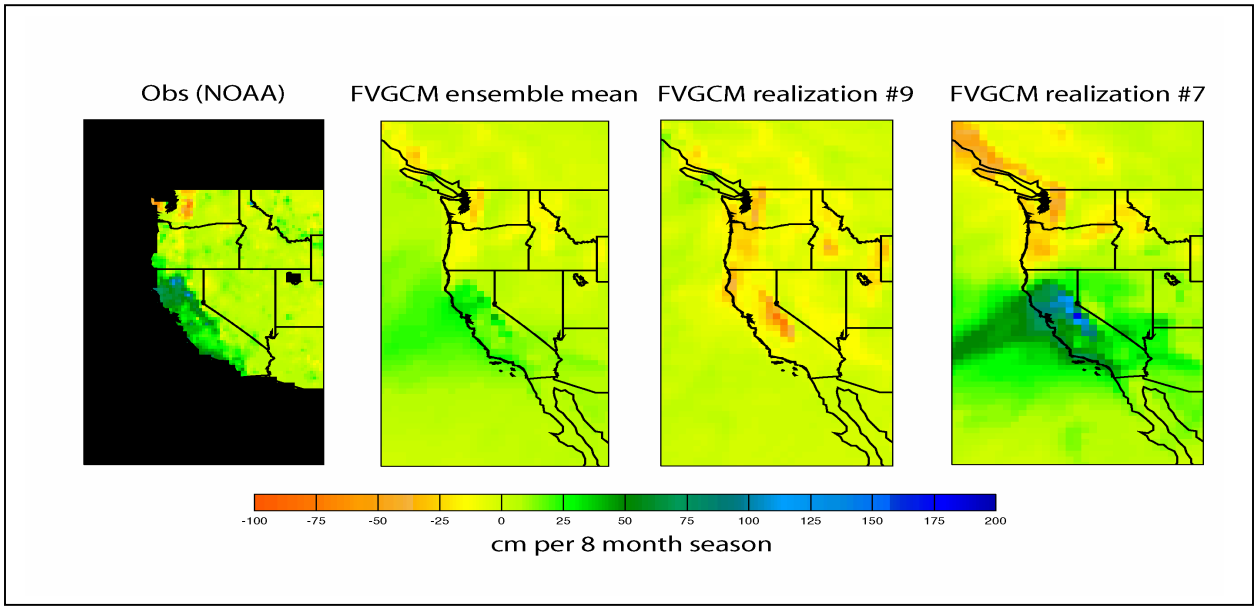
A1.10 High-Resolution Nested-Model Simulations of Annual Mean Precipitation

The figure shows annual mean precipitation in California—Nevada simulated by a regional (limited-domain) climate model at 9 km resolution nested within a global climate model at 75 km resolution. This nesting approach is widely used, but the grid dimensions here are about four times smaller than typical. A comparison to an observation-based estimate of precipitation (left panel) shows that the simulated precipitation has realistic spatial structure and minimal large-scale biases. The realistic spatial detail results from the high resolution of the nested model, which allows realistic representation of topographic variations. The minimal large-scale biases results from the high resolution of the driving global model.
— Phil Duffy



A1.11 Predictability of Precipitation in California during El Niño Winters

We have been investigating the theoretical predictability of precipitation in California associated with El Niño. We performed ensembles of simulations on Thunder of different El Niño winters (1991–1992, 1997–1998, and others). These used a global climate model at 0.5 deg (latitude × longitude) resolution. This is very fine resolution for global climate simulations. For each winter we performed 10 to 12 simulations using the same boundary conditions (observed sea-surface temperatures) but different initial conditions. Each simulation covers eight simulated months. Comparing the different ensemble members allows us to assess the relative influence of boundary conditions (sea-surface temperatures) vs. initial conditions on precipitation. Strong agreement among the ensemble members would indicate strong influence of sea-surface temperatures on precipitation and thus would indicate that precipitation is *in theory* predictable. Weak agreement among ensemble members would indicate strong influence of initial conditions. Because the atmosphere is chaotic, this would mean that precipitation is not predictable, even in principle. In the figure below, the left panel is observed precipitation anomalies (departures from average) for the winter of 1997–1998 in the Western U.S. It was very wet in California and very dry in the Pacific Northwest. This pattern is typical of El Niño, but the amplitude is unusually strong. The second panel shows the mean precipitation anomaly from an ensemble of 12 simulations performed on Thunder. (As indicated above, these all used prescribed, observed sea-surface temperatures as a boundary condition, but each had different initial conditions). These ensemble-mean precipitation anomalies agree pretty well with observed anomalies. (wet in CA, dry in Pacific Northwest.) The third and fourth panels show results of individual simulations from the ensemble; the third represents a severe drought and the fourth is tremendously wet. This means that when we have an El Niño like the one in 1997–1998, the most likely result is close to what actually happened (wet in CA, dry in Pacific Northwest) but there is some probability (on the order of 10%) that it will be incredibly wet in California and about the same probability that we will have a drought. The best we can do to predict this is to make a probabilistic forecast.
—Phil Duffy



Appendix 2. Detailed Usage Requirements

Below are the actual requirements as they were collected from the scientists. M&IC provided the scientists with a template showing requirements collected from the survey in 2001 and then asked them to respond by indicating any new or updated requirements. For the purposes of this survey, capability requirements were considered to represent calculations that would require access to the proposed 40 TF Peloton. Capacity requirements may be thought of as representing day-to-day needs for runs in much greater quantity but at much smaller scale. This would include the necessary small-scale runs to develop intuition required to prepare for meaningful capability runs.

	Project ID	Capability Requirements (40 TF)		Disk	Archive	When Ready?
		Time Needed	Mem/Node			
1	BioImaging	injection: 9 days/year damage: 9 days/year classification: 11 days/year	injection: 2.5 GB/node damage: 16 GB/node classification: 8 GB/node	injection: 6 GB damage: 6 GB classification: 1 TB	injection: 0.4 TB damage: 0.4 TB classification: 4 TB	Jun-06
2	Z3	~ 15 days per run, multiple runs (1 per month)	>~2 GB per processor	4-6 TB per run	~100 TB per run	Now
3	BOUT	100 runs, ~20hrs /run	1.5 GB per processor	<~1TB	<~ 100 TB	Now
4	QCD/PHENIX					
5	pF3d	1-4 mo /calc -- 1-2 major calcs (over next year)	3-6 TB total => ~1-2 GB /CPU	~20 TB (restarts and history)	~160 TB	Now
6	HP-CFD	1/2 machine every couple of months, 2 wks/run	>2 GB per node			May-02
7	Climate	several simulations per year at 10^6 P-hr	2-4 GB /node?	1 TB local disk	5 TB	Now
8	AtmosChem	2 studies @ 200 nodes @2wk	0.5 B/F => 16 GB /node	5-10 TB	5 TB	Now, for some studies - largest simulations ready ~9/05
9	Homeland Security (Structures)					
10	GEODYN/ RAPTOR	500,000 CPU-hours/run; one run every few months	8 GB per node	~ 10 TB per run	~10 TB per run	When scaling issues (currently being investigated) are resolved
11	LDEC	multiple runs, 8 wk total	>~1GB per processor	1000s GB per run	several TB	Now
12	GEODYN/L	not available, scaling & timing studies needed	8 GB per node	~ 10 TB per run	~10 TB per run	parallel version under development
13	HighPlowZ	3 ongoing projects that will require ~12 days/simulation on 8 TF machine	2 GB/proc	~1 TB on gpfs	~10 TB	Now
14	Compnano	12 days/run, with 5 runs/year	8 GB/node (4 proc/node)	1.5 TB	10-15 TB	Now
15	Beryllium	No capability requirements for current project goals				

	Project ID	Capability Requirements (40 TF)		Disk	Archive	When Ready?
		Time Needed	Mem/Node			
16	Nanosim	Ongoing Grand Challenge allocation on Thunder. Current allocation is ~400K CPU hours/week	8 GB/node	5 TB on parallel filesystem	20 TB on storage	Now
17	EvRec	60,000 CPU h/month	2 GB /node	0.5 TB	3 TB	Now
18	Tempest	1 week / run in 5D, 10 runs per year	2 TB / # of nodes	60 TB	100 TB	Apr-06
19	Folding	8 simulations each run on 2 BG/L racks (16 BG/L racks for 10 months)	<0.25 GB/node	100 GB	1 TB	Now
20	MDHydro	RT: 50K CPU-hrs/week, RM: 25K CPU-hrs/week	> 512 MB/proc	>10 TB	>10 TB	Now
21	Rad Effects	Rad delta: 10K CPU-hrs/week, Rad alpha: 2K CPU-hrs/week, L3 Milestones	> 512 MB/proc	1 TB	1 TB	Sep, 2005
22	Aging Effects	Isentropic 20K CPU-hrs/week, Shock 20K CPU-hrs/week, L3 Milestones	>1 GB/proc	5 TB	several TB	Now
23	VASP	phase stability: 10k CPU-hrs/week, Defects: 5k CPU-hrs/week	>1 GB/proc	250 GB	1 TB	Now
24	mesochem	1 month for 1 project - seveal projects are ready Now - limited only by CPU time avaliable	1 GB per node probably sufficient.	8 TB per run	several TB	Now
25	ParaDiS	~20 runs / 1-week long each run	<1 GB node	200 Mb per run	20 Gb	Now
26	Qchem	48h-74h/run 100 jobs, say, 480-600h over 12 months	2-4 GB /node	0.5-1 TB	<~1 TB	Now
27	PIMCQVIB	200/h runs, multiple runs	100 MB to 1GB per node	1GB per run	250 GB	Now
28	ShockRad	40 runs, 1 day/run, about 300,000 CPU-hours	<1 GB/node	1 TB		Now
29	Poly	PDMS simulations: 1/3 year@ 2048 processors; Polymer crystalization:1/4 year@ 2048 processors; Data analysis 100h @1024processors every week	2-4 GB per node	1 TB	10 TB	Now
30	Computational Chemistry and Biology	20% of machine	< 500 Mb			Now
31	Superionic	Multiple 1 wk runs, several months, total	min 1 GB /node – 2 GB /node better	>20 GB		Now
32	ShockMetal	~50 runs with 2.5K-100K CPU hours/run	0.2-2 GB/node	50 TB	100 TB	Now
33	astroMat	~50 runs with 2.5K-100K CPU hours/run	0.2-2 GB/node	50 TB	100 TB	Now

	Project ID	Capacity Requirements		Disk	Archive	When Ready?
		Time Needed	Mem/Node			
1	BioImaging	injection: 65 days/year damage: 90 days/year classification: 125 days/year	injection: 250 MB/node damage: 2.5 GB/node classification: 8 GB/node	injection: 6 TB damage: 6 GB classification: 0.5 TB	injection: 12.5 TB damage: 0.4 TB classification: 1 TB	Jun-06
2	Z3	3-7 days per run, many capacity runs per each capability run	>~1 GB per processor	~0.25 TB per run	~5 TB per run	Now
3	BOUT	50 runs, ~50hrs /run	1.5 GB per processor	<~1 TB	<~ 100 TB	
4	QCD/PHENIX	100/200 GF	4 GB		10 TB	
5	pF3d	2-3wk @128-256P/calc (300-600 GF/s) x ~15 calc	64 GB total => 0.5 GB /CPU	0.5 TB	~20 TB	Now
6	HP-CFD	2 runs/wk, 2days/run	>2 GB per node			Now
7	Climate	twenty calculations per year at 5*10^5 P-hr	2-4 GB/node	0.5 TB	2 TB	Now
8	AtmosChem	1-2 wk/run, multiple studies	4 GB /4P node	5 TB	10 TB	Now
9	Homeland Security (Structures)	5 multi-day runs weekly	2 GB /node	100 GB/per simulation	5 TB	Now
10	GEODYN/ RAPTOR	~10,000 CPU-hours per week	1-4 GB per node	100's GB per run	100's GB per run	Now
11	LDEC	100s of parameter study runs using 5-25 days on 8-256 CPUs	>~1 GB per processor	10's GB per run	several 100 GB	Now
12	GEODYN/L	not available, scaling & timing studies needed	1-4 GB per node	100's GB per run	100's GB per run	parallel version under development
13	HighPlowZ	Each of the 3 major projects requires ~5 4-12 hour jobs on an ongoing basis.	~1 GB/proc	10 GB per project	5 TB	Now
14	Compnano	Many 6-12 hour jobs	8 GB/node	5 GB/job	5 TB	Now
15	Beryllium	8, 6.3K CPU hour simulations (on 33 Thunder nodes) and 16, 18K CPU hour jobs	4 GB/node	10 GB/job	10 TB	Now
16	Nanosim	We have approximately 6000, 30 node, 2-hour jobs to run	4 GB/node	50 GB	10 TB	Now
17	EvRec	20,000 CPU h/month	2 GB/node	0.5TB	1 TB	Now
18	Tempest	12 days / run in 4D, multiple studies	100 GB / # of nodes	3 TB	10 TB	Sep-05
19	Folding	256 simulations run on a total of 12 BG/L racks for 1 month	<0.25 GB/node	100 GB	1 TB	Now
20	MDHydro	RT: 20K CPU-hrs/week, RM: 20K CPU-hrs/week	> 512 MB/proc	1 TB	1 TB	Now
21	Rad Effects	Rad delta: 20K CPU-hrs/week, Rad alpha: 5K CPU-hrs/week	> 512 MB/proc	1 TB	1 TB	Sep. 2005
22	Aging Effects	Isentropic 5K CPU-hrs/week, Shock 5K CPU-hrs/week, L3 Milestones	>1 GB/proc	5 TB	several TB	Now
23	VASP	phase stability: 10k CPU-hrs/week, Defects: 5k CPU-hrs/week	>1 GB/proc	250 GB	1 TB	Now

	Project ID	Capacity Requirements		Disk	Archive	When Ready?
		Time Needed	Mem/Node			
24	mesochem	100s of 12-hour runs	~1 GB per processor (like Thunder)	1 TB		Now
25	ParaDiS	200 runs / 3 days each run	< 1 Gb node	40 Gb	40 Gb	Now
26	Qchem	Job needs some times up to one week of CPU time on a single node with several processors	2-6 GB /node	1-2 TB global disk	<~1 TB	Now
27	PIMCQVIB	not available, tweaking of code and scaling studies needed				assuming in 12 months
28	ShockRad					
29	Poly	PDMS simulations: 2/3 year@ 1024 processors; Polymer crystalization:1/2 year@ 1024 processors; Data analysis 200h @512processors every week	1-2 GB per node	1 TB	10 TB	Now
30	Computational Chemistry and Biology	15% machine	< 500 Mb			Now
31	Superionic	100s of 12 hour runs	1 GB/node	10s of GB per run		Now
32	ShockMetal	4 runs, 300K CPU hours/run	0.2-2 GB/node	100TB	200 TB	Now
33	astroMat	4 runs, 300K CPU hours/run	0.2-2 GB/node	100TB	200 TB	Now