SAND2007-4516P

BACKGROUND INFORMATION AND TECHNICAL OVERVIEW OF CCIM

CIS External Technical Review, July 25-26, 2007
James Peery, Director 1400
Computation, Computers, Information, and Mathematics Center
Sandia National Laboratories
Sand Report xyz

Computer and Informational Sciences (CIS) Council:
The Vice President for Science, Technology & Engineering (CTO, Division 1000) manages science, technology, and engineering (ST&E) R&D at SNL through a set of 8 ST&E Councils: Biotechnology; Chemical and Earth Sciences; Computer & Information Sciences; Micro-electronics & Photonics; Engineering Sciences; Manufacturing Sciences; Materials Science and Technology; Pulsed Power. Each of these councils administers R&D programs, funded both internally and externally.

The ST&E Councils create new science, engineering, and technology that enable Sandia's four other programmatic Management Units to achieve SNL's national security missions. Sandia's six Strategic Management Units (SMUs) fall into 3 groups: The Nuclear Weapons SMU constitutes the first group; the Energy, Resources and Nonproliferation, Homeland Security and Defense, and Defense Systems and Assessments SMUs constitute the "Integrated Technologies and Systems" group; and the ST&E and Integrated Enabling Services SMUs constitute the "Laboratory Transformation" group.

The CIS Council comprises two Centers: Computation, Computers, Information, and Mathematics (CCIM, 1400) and Computer Sciences and Information Technology (CSIT, 8900). The council members are the Directors of the two Centers, James Peery (1400) and Len Napolitano (8900), and the Senior Managers responsible for R&D in these two Centers: Howard Hirano (Computer Sciences, 8960), Jennifer Nelson (Computer Science R&D, 1430), Sudip Dosanjh (Computer and Software Systems, 1420), David E. Womble (Computer Sciences & Mathematics, 1410). Research programs administered by the CIS Council or its constituent Centers include the Enabling Predictive Simulation LDRD area (managed jointly with Engineering Sciences), the Computer Science Research Foundation (CSRF), which is funded through the NNSA Advanced Simulation and Computing program (ASC), and the ASC-funded Computer Sciences Research Institute (CSRI), which is dedicated to interactions and collaborations with academia by funding collaborative visits by students and faculty, post doctoral appointments at Sandia, and workshops.

CCIM Vision:
The Center for Computation, Computers, Information, and Mathematics (CCIM) will provide national leadership in high performance computing for science and engineering. We will establish the Center as a national leader in informatics. We will ensure our R&D capabilities are applied to the critical missions of the laboratory.

CCIM Mission: Enabling High Performance Computing. More specifically, the CCIM mission is

1. To provide SNL with the computational and informational science tools and the platforms they need to carry out the labs science and engineering missions at the leading edge compared to our competitors;

2. To be a seed for new lab science initiatives: *e.g.* in informatics and computing for intelligence, in applying cognitive science to our missions, and in building modeling and simulation capabilities supporting nanoscience and global nuclear energy initiatives
3. To help disseminate new CIS technologies throughout the lab;
4. To be a high-tech corridor from academia into Sandia and from Sandia into industry (spin-in and spin-out).
5. To be a national resource in computing and informatics for America's national security missions.

Our future lies in fundamentally new approaches to computing that will be needed in about a decade (the post peta-scale era): stochastically available computing systems and robust self-adapting and self-healing applications. It also lies in information science and informatics, including bioinformatics and in the marriage of informatics and simulation science.

CCIM is presently conducting strategic planning exercises and has completed drafts of 1) strengths, weaknesses, opportunities and threats and 2) mission and vision statements. Discussion on thrust areas is ongoing; however, it is clear that a base competency of scalability and a core theme of transformational applications are emerging. A <u>new</u> CIS strategy has been developed in parallel and is provided below. As one would expect there are synergies between these two strategies. The CIS strategy will be discussed with the review panel on the second day of the review.

<u>CIS Vision:</u> Provide national leadership in high-performance computing computer science and mathematics; develop enabling technologies and solutions to our nation's most challenging national security problems.

<u>CIS Mission and Scope:</u> The Computer and Information Science Research Foundation will develop and maintain a high-performance computing capability at Sandia National Laboratories, as well as the expertise to apply those capabilities to a wide range of national security problems.

The research foundation will work towards three objectives:

- Deliver computer science, mathematical science, computational science and information science tools and platforms that enable Sandia's mission.
- Be national leaders in high-performance computer architecture, system software, algorithms, transformational technologies, and modeling and simulation in critical application areas. Drive transformation through predictive simulation.
- Anticipate future national security needs in computer science, mathematical science, computational science and information science and be a catalyst for new initiatives.

We will establish metrics to measure successes and progress towards these goals.

Charge to the Review Panel:
The role of the CIS Review Panel is both to assess the work being done by CIS organizations and to advise these organizations on current and future directions. In 2007, the review panel is expected to focus on the advisory aspect with specific consideration of the CIS strategy and strategic plan. In 2008, the review panel will assess the CIS R&D program as part of the Laboratory Performance Evaluation and Assessment report, which is part of the Sandia Corporation contract fee determination from NNSA The Review Panel members have been selected by the University of Texas in order to insure independence of the review process. Each panel member has a responsibility to contribute independently and without personal or institutional bias.

We ask that you measure the quality and appropriateness of the CIS Council's S&T R&D programs - to the extent possible within the context of Sandia's missions and in the context of the national R&D enterprise.

We want your evaluation of our areas of strength and our weaknesses, as well as a holistic assessment of our place in the international community in high-performance computing and technical informatics. The metrics by which we should be evaluated include

- Quality of work;
- Quantity of work;
- Value for the dollar;
- Innovation, creativity, and originality;
- Successful application of our R&D to Sandia, DOE, other U.S. Government Agencies (*e.g.*, DoD, DHS, NSA, EPA), and industrial needs;
- Peer valuation of the work measured by publications and citations to them.
- Recognition of leadership in high performance computing through awards, invited talks and papers, and key committee membership/leadership roles.
- Effectiveness in utilizing research funding to set a foundation for the future.

Specific to the new strategic plan, we ask that you advise us on

- National and international trends
- Opportunities that should be explored
- Culture and environment (i.e. innovation, creativity, collaboration, etc.)
- Strengths and weaknesses

The full scope of our R&D pursuits is too extensive to be well represented and still keep the duration of the review reasonable. Accordingly, we select presentations each year with the aim of covering the full scope of CIS activities over ***three*** successive reviews.

Customers for this review:

1. The principal customer is Dr. Rick Stulen, the Vice President for Science, Technology, and Engineering / Chief Technology Officer at Sandia.
2. The second customer is the management team in CIS, which relies on you to provide annual feedback on our progress and suggestions for new directions. We also rely on you to help validate some of the issues that we are bringing to the attention of Sandia and DOE management.
3. The third customer is the U.S. DOE, which relies on the reports from our external reviews in assessing the quality and success of our R&D efforts.

A Challenging Year with Significant Successes:

CCIM continued to excel technically this past year. Sandia's partnership with Cray helped revitalize the supercomputing industry. Cray won many contracts based on the Sandia/Cray Red Storm architecture. There are now 18 worldwide Cray XT sites and over 25 systems have been deployed. This was an impressive turn-around for a company that was close to bankruptcy a few years ago. Sandia upgraded Red Storm, adding a 5th row and replacing all of the processors with dual core opterons. This boosted Red Storm to second place on the top 500 list (101.4 Tflops) when it was released at the SC06 conference. More importantly this upgrade has resulted in significantly improved application performance, close to 70% speedup on many problems. Last year John Daly (LANL) gave a customer perspective presentation at our external review. John ran a 5000-processor simulation for many months. After the upgrade, he continued running his job on 3000 dual core processors with the same performance.

Other highlights included winning an R&D 100 Award for the computer processor allocator, best paper award at European Supercomputing and best paper award at the International Parallel and Distributed Computing Symposium. Mark Boslough's asteroid impact study on Red Storm was featured in a BBC and National Geographic documentary on the formation of Libyan Desert Glass and was named one of the Top 100 science stories of 2006 by Discover magazine. Paul Taylor's work on traumatic brain injuries was featured in Medical Imaging Magazine. And finally a map of science developed by Sandia researchers was published in Nature and displayed at the New York Hall of Science.

Despite these and other technical successes, Sandia faced significant programmatic challenges. Sandia received a disproportionate cut to its ASC program and NNSA indicated that capability computing would be consolidated to two sites. Although NNSA has not officially named the two sites, FY08 budget targets suggest that the two sites are LLNL and LANL. During this time Bill Camp announced his retirement and left the laboratory to become the Intel's Chief Technology Officer for High Performance Computing. Sudip Dosanjh served as acting director from December 2006 to April 2007 when James Peery was named the new CCIM director. Sandia management (Tom Hunter, Rick Stulen, Tom Bickel, James Peery, Art Hale and Sudip Dosanjh) has worked hard to try to reverse this apparent headquarter's decision. These actions along with advocacy from external High Performance Computing (HPC) experts has at this time helped NNSA understand the important national role SNL has had in architecting and engineering balanced HPC capability platforms. NNSA has indicated that they expect SNL to continue in this role but the details required to ensure resource retention at SNL has not been completed. We have also aggressively pursued other sources of funding and impact including:

- In collaboration with Pittsburgh Supercomputing, Intel and Hewlett-Packard, we submitted a proposal to NSF to build and site a multi-petaflops supercomputer.
- We received Office of Science funding to develop a quad-core version of the Catamount light-weight operating system and we are negotiating a work for others agreement with Intel to extend this work to many core architectures.
- We are working to establish an Institute for Advanced Architectures as suggested by last year's panel and now called out in the Senate version of the FY08 Energy and Water appropriations bill. This includes establishing closer ties to Office of Science,

NSA and DARPA and building collaborations with Intel, Sun, Cray, SGI, IBM and Micron.

- We are partnering with Governor Richardson's office, LANL and New Mexico universities to create a state supercomputing center.
- We have proposed to ASC-HQ another upgrade of Red Storm.

There has also been a major change in facilities occupied by CCIM staff; roughly half our staff are now in the Computer Science Research Institute (CSRI). The 34,500 sq. ft. building includes a large number of visitor offices and shared work areas, and is an outstanding facility for collaborative research between Sandia scientists and university and industry researchers.

Finally, we note some new directions for which we are pursuing or preparing:

- We are increasing our strategic focus in the area of informatics, which is broadly defined to include computing for systems and data that are not geometric in nature. This effort crosscuts our computing capability and includes work in computer architectures, problem setup, databases and data representation, algorithms and code development, visualization, analysis and application. The applications of our informatics work are far reaching, including biology, complex systems, decision analysis and national security (intelligence).
- In response to President Bush's call for renewed development of nuclear energy, under his Global Nuclear Energy Partnership (GNEP) program, CCIM is actively supporting the Energy, Resources, and Nonproliferation SMU in engaging the multiple National Laboratory group that is working to define a technically viable program. We have collected our staff member's considerable experience in nuclear reactor safety modeling from their pre-CCIM days and are combining that with our expertise in high performance computing and Sandia's experience and products from its ASC efforts. It will be highly challenging, technically, to 'close' the nuclear fuel cycle while maintaining both safety and security.
- Partnering with the Defense Systems and Assessments SMU, we are increasing our collaborations with DoD in both physics based and decision based applications.

Highlights of Past Year:

Much has been accomplished in the past year. Some of our highlights are given below. These and others will be discussed in more detail during the technical presentations.

Compute Process Allocator Receives R&D100 Award: The Compute Process Allocator (CPA), led by Vitus Leung, received a 2006 R&D 100 Award and a Sandia National Laboratories' team Employee Recognition Award. The CPA assigns parallel processing resources on supercomputers. CPA is the first allocator to balance individual job allocation with future allocation over 10,000 processors, allowing jobs to be processed faster and more efficiently. In experiments, the algorithms in CPA increased the locality and throughput on a parallel computer by 23% over simpler allocators.

A Mathematical Analysis of AtC Coupling Methods: We are finalizing the paper "On Atomistic-to-continuum (AtC) Coupling by Blending" that presents a mathematical framework for blending atomistic and continuum models over a physical subdomain. This builds upon the work presented in the two papers and "Concurrent AtC Coupling based on a blend of the Continuum Stress and Atomistic Force" and "A Force-Based Blending Model for Atomistic-to-Continuum Coupling" accepted for publication in Computer methods in applied mechanics and engineering, and the International Journal for Multiscale Computational Engineering, respectively.

Advanced Modeling and Decomposition Techniques for Large-Scale Combinatorial Problems: Many important problems in scheduling and resource allocation, for example, can be modeled and solved using integer programming. A critical measure of the hardness of an integer program is whether there is a provably good linear programming approximation (a finite integrality gap). Sandia has developed the first algorithm that will take such an approximation and find an integer solution whose quality is determined by this gap. We have recently incorporated this capability into PICO, along with a "feasibility pump" heuristic that performs well in practice.

2007 International Parallel and Distributed Processing Symposium (IPDPS) Best Paper Award: Researchers from Sandia National Laboratories and the Ohio State University were awarded the Best Paper award in the Algorithms track of the 2007 International Parallel and Distributed Processing Symposium (IPDPS). Their paper, "Dynamic Load Balancing for Adaptive Scientific Computations via Hypergraph Repartitioning," was judged best of the 419 submissions to the Algorithms track. It presents a novel algorithm for redistributing data in adaptive parallel simulations to maintain processor load balance as computational requirements change. The new algorithm exploits a robust and accurate hypergraph-partitioning model to reduce average total communication costs by roughly 20% compared to traditional graph repartitioning methods.

Solver Improvements for High Energy Density Physics (HEDP) simulations: Chris Siefert, Ray Tuminaro, and Jonathan Hu have completed development of a discrete reformulation of, and a new multilevel solver for, the eddy current Maxwell's equations for edge element discretization. The multilevel solver builds on existing solver technology for vector nodal Laplacians, creating the potential to apply advanced techniques for those problems to Maxwell's equations with the potential for direct impact to the HEDP focus area. Preliminary results are very promising, demonstrating improved convergence with the possibility of improved parallel scaling as well. A journal publication is in preparation and the algorithm is being deployed within the Trilinos ML package.

Automatic Differentiation Capabilities Deployed: The Sacado Automatic Differentiation (AD) capability is being deployed in the Aria application code, enabling users to essentially define whole new equations with full correct, analytical (FAD) sensitivities via a user plug in. Sacado is being funded under the ASC Computer Science Research Foundations (CSRF) to develop analytic derivative capabilities in Sandia simulation codes. A benefit for Aria users is that planned extensions to Sacado will automatically generate uncertainty quantification (UQ) information without additional development.

Progress Towards An Integrated QMU Capability: We have successfully applied large-scale nonlinear optimization algorithms in MOOCHO to solve large-scale inverse problems modeled in Charon in parallel. Early results show that the inverse solve scales the same as the forward solve while varying the numbers of processors and overall parallel scalability is only limited to the performance of the preconditioned iterative linear solver (which is the case for the forward implicit solver as well). This inverse problem is a critical stepping-stone to our Level 2 Milestone's targeted goal of automatic solving parameter estimation problems against experimental data. Charon is Sandia's device modeling code, and the L2 milestone is to provide an integrated solver/inversion capability for Sandia's quantified margins of uncertainty (QMU) efforts.

The CSRI Moves Into New Facilities: CSRI's first year in our new building has highlighted the benefits of staff-collocation and dedicated visitor and collaboration space. In 2007 CSRI as hosted 72 short-term visitors, 11 long-term faculty visitors, 4 CSGF fellows, 32 summer students and 6 year-round graduate students. CSRI has launched a new web site (csri.sandia.gov) and has published its first Summer Proceedings (csri.sandia.gov/Proceedings). In 2007, CSRI will host four technical workshops (two offsite and two in our new CSRI facility) including a groundbreaking workshop on Mathematical Methods for Verification and Validation being organized by our 2007 von Neumann Fellow, Clayton Webster.

Charon Used to Design Air Force SOG reactor: The Air Force Research Laboratory recently selected researchers in CCIM to lead the development of a parallel computational simulation and design capability for the singlet-oxygen generator (SOG). The SOG is the initial stage in the chemical oxygen iodine laser (COIL), which has important applications for both military purposes and for industrial cutting and drilling. A multi-disciplinary team from CCIM and Engineering Sciences is creating a three-dimensional multiphase reacting flow model within the finite-element research code Charon to predict flow, species conversion and resulting efficiencies, utilizations, and yields for various configurations of the SOG reactor.

Dakota/Xyce Used to Characterize Manufacturing Uncertainties: CCIM's software and expertise enabled a cross-center multi-disciplinary team to conduct crucial pre-fabrication performance and reliability assessments of an emerging CMOS7 integrated circuit with potential application to weapons systems. Center experts in electrical circuit modeling/simulation and optimization/uncertainty quantification teamed with device physicists, electrical engineers, and circuit designers (from Centers1700, 8200, and 1300) to perform the first simulation, sensitivity analysis, and uncertainty quantification for Sandia's CMOS7 ViArray. The electrical circuit code Xyce provided efficient parallel simulation of circuit behavior. Coupled with Xyce, DAKOTA was used to perform iterative analyses to assess the effect of manufacturing variability and uncertainty on device response range, and to determine which process layers most significantly influence key performance metrics.

Improved Logistics Optimization Capabilities: The Optimization team for enterprise logistics has implemented a specialized simulation framework for representing logistics operations, where the value of our improved solutions can exceed $100M. Simulations developed with our framework are 100 to 1000 times faster than simulations of comparable detail implemented using commercial simulation tools, enabling us to find optimal solutions to large-scale logistics operations, such as the Joint Strike Fighter.

Water Security Toolkit Developed: As part of Sandia's ongoing collaboration with the National Homeland Security Research Center of the Environmental Protection Agency (EPA), Sandia has developed a software package called the "TEVA-SPOT Toolkit." This encapsulates several years of algorithmic research on the problem of placing water quality monitoring sensors into drinking water distribution networks. The software is now a key product of the EPA's Threat Ensemble Vulnerability Analysis (TEVA) program, which also includes collaborators from Argonne National Laboratory. The TEVA-SPOT Toolkit offers solvers that can handle problems of much greater size than current commercial tools, and also can provide confidence bounds regarding the optimality of its solutions. Furthermore, TEVA-SPOT solvers can optimize multi-objective problems using a novel mathematical programming approach. These capabilities are at the forefront of the water security field.

New Visual Analytic Techniques Developed: Within our CSRF Decision-Support Informatics research, new visual analytics techniques were developed and prototyped: state-of-the-art optimization/sampling methods, nonlinear dimensionality reduction techniques, and visualization capabilities for interactive examination of complex, high-dimension search spaces to gain deeper insight into model structure and search-space tradeoffs and to accelerate global optimization and analysis. These techniques have been integrated into a common framework with a prototype demonstration this quarter on a polypeptide folding problem, chosen for its challenging degree of complexity.

Trilinos Lifecycle Model Developed: The Trilinos Lifecycle Model team has recently had a paper entitled "The Trilinos Lifecycle Model" by J. Willenbring, M. Heroux and R. Heaphy accepted for presentation and inclusion in the IEEE conference proceedings of the SE-HPC 2007 Workshop on software engineering for high performance computing as part of the 29th International Conference on Software Engineering, being held May 2007 in Minneapolis, MN. The Trilinos Lifecycle Model described in this paper is a three-phase promotional model intended to support software engineering through: basic research, production-growth, and production-maintenance, with well-defined events that are required for promotion from one phase to the next.

RedStorm Upgraded to 125TF: Sandia and Cray successfully upgraded Red Storm by adding a 5th row, exchanging the single-core 2.0 GHz AMD Opterons for dual-core 2.4 GHz processors, doubling the interconnect performance, and upgrading Sandia's Catamount system software. These upgrades increased Red Storm's measured HPL/Linpack speed from 36.19 Tflops to 101.4 Tflops. When the Top500 supercomputer list was announced at Supercomputing 2006, Red Storm placed second overall, and first among the general-purpose high-end computer systems that are useful for the broad range of our engineering and scientific analysis codes.

Advances in Unconstrained Plastering: Despite continuous research for more than two decades, no algorithm has been developed which robustly generates conformal, high-quality, all-hexahedral finite element meshes on arbitrary geometry assemblies. Numerous algorithms have been presented, although none meet all the strict requirements of:

1. All-hexahedral elements,
2. Minimized distortion and skew, especially near the boundary,
3. Conformal elements between components in general assemblies,
4. Conformal elements through out each volume (i.e. no hanging nodes).

However, a new algorithm called Unconstrained Plastering is showing great promise. Unconstrained Plastering is based upon geometry decomposition, which is guided by advancing fronts from an unmeshed volume boundary. Geometry decomposition continues until each of the partitions can be meshed individually with sweeping, mapping, or midpoint subdivision. Progress and implementation on Unconstrained Plastering has matured enough to successfully generate high quality meshes on more complex solid models as illustrated in the figures below. Models with concavities and curved surfaces are handled. Issues with high valency internal vertices have been addressed successfully. Research is currently focused on handling complex front collisions resulting from more complicated geometries, generating meshes on assemblies, and the recombination of similarly constrained adjacent partitions.

Immersive Topology Engine for Meshing: As part of an ASC level 2 milestone, the CUBIT team has developed a new environment for guiding analysts through the process of generating a hexahedral or tetrahedral mesh for simulation. This new environment is called the Immersive Topology Engine for Meshing (ITEM) and is built on the existing CUBIT Geometry and Meshing Toolkit. New geometric reasoning algorithms have been developed that can detect potential problems in a CAD model and provide a list of suggested solutions. This is offered in a wizard-like environment where the user may systematically step through the geometry preparation and meshing process and run a series of diagnostic tests. The user is then presented with a list of solutions to specific geometric problems that can be easily previewed and performed at will. With phase 1 of ITEM now complete, this tool can now generate diagnostics and solutions for a range of geometric problems including resolving small features, detecting and fixing imprint/merge problems and detecting and suggesting potential decomposition options for hexahedral sweeping. Scheduled to be completed August 2007, this new tool promises to dramatically improve the productivity of analysts who currently must spend considerable time developing meshes for simulation.

Scalable Information Visualization: The Data Analysis and Visualization department, has released a beta version of its scalable information analysis tool ThreatView™, and is now working directly with analysts in the intelligence community to refine the tool in preparation for its deployment and use in a production environment. ThreatView is designed to enable analysts to explore, understand, and manipulate abstract, complex data such as vast collections of communications (email, etc.), documents (papers, patents, etc.) or other geospatial data. ThreatView's scalable architecture significantly expands the size and complexity of data that customers can analyze – the tool can run on anything from a laptop to a supercomputer – and is designed to work with everything from large databases to legacy collections of documents. In addition, ThreatView provides a broad palette of views – ways of visualizing data – that include landscape views (documents are grouped into 'islands' by similarity), geospatial views (maps, globes, etc.), and detailed information views for document, attribute, and metadata tables. In addition, ThreatView enables link analysis, allowing analysts to drill into specific information about links between the items they are viewing. Funded in part by National Ground Information Center's (NCIG) PATTON Alliance, ThreatView continues Sandia's innovation in information visualization by leveraging Sandia's strengths in large-scale data visualization

CCIM Restructures To Strengthen Material Dynamic Response Area: The core technical direction of CCIM's computational materials efforts was refocused by moving the staff working on Peridynamics/EMU team to 1435. This addition gives us a secure foothold of expertise at the continuum scale and Peridynamics fits with our on-going focus on simulation methods development for materials investigations. The interests and expertise of the staff, collectively, now give the department a significant capability in modeling the dynamic response of materials, including the chemistry and reactivity that accompanies or drives dynamic response. As this area is of basic interest to Sandia, all the NNSA labs, and to DoD, we will emphasize this capability of the department and strengthen it. To this end, the department name was changed to "Multiscale Dynamic Material Modeling"; previously our name was "Multiscale Computational Materials Methods".

CCIM Gains A Truman Fellow: O. Anatole von Lilienfeld-Toal was selected by Rick Stulen (VP Science & Technology and CTO) as one of two new FY2008 Truman Fellows, Sandia's most prestigious position for recent Ph.D.s. Anatole has joined CCIM to pursue his three-year long research project entitled "Multiscale schemes for the predictive description and virtual engineering of materials." Ann Mattsson and Peter Schultz are mentoring him, but he will be collaborating with Sandia staff in several departments throughout SNL, as well as in the external research community.

Advancements Made in Peridynamics Mathematics: Stewart Silling and Rich Lehoucq derived an expression for a stress tensor consistent with Silling's "Peridynamics" reformulation of continuum mechanics. Although stress has no fundamental role in Peridynamics, this new peridynamic stress tensor could provide a way of directly connecting an atomistic description of a material with a continuum model. They are now exploring how this may enable bridging the length scales within a single consistent mathematical model. The paper "Force Flux and the Peridynamic Stress Tensor" has been submitted to the Journal of the Mechanics and Physics of Solids.

MHD Modeling Continues to Have Major Impacts: The ALEGRA shock and multiphysics application code, under development within the center for many years under the ASC program, has been applied recently in several areas with great success. As we reported last year, the MHD capabilities in ALEGRA were used to help design a series of magnetic flyer plate experiments on the Z machine to characterize the beryllium shock melt regime. These modeling and simulation efforts resulted in the savings of months of experimental time and hundreds of thousands of dollars in the reduced (by half) number of Z shots. Similar savings were achieved by application of this predictive capability in subsequent dynamic materials experiments for high-density carbon and plutonium, and should result in enormous savings with the upgraded ZR machine when it returns to service. ALEGRA simulations of wire array experiments on ZR are now being used to understand the dynamics of z-pinch implosions, which are critical for optimizing the radiation source for science-based stockpile experiments. ALEGRA is also being used quite effectively to support the electromagnetic armor program at the U.S. Army Research Laboratory. Recent simulations have for the first time captured major features observed experimentally but which are not well understood, providing new insights to the experimental program.

Version 1.0 of the Omega-AB code has been released: The initial design of Omega-AB was intended to provide a purely agent-based simulation environment that was GUI driven. Version 1.0 was designed so that Omega-AB can also support general discrete-event, systems dynamics, and hybrid modeling paradigms. The ability to support hybrid simulations is extremely useful in that complex environments typically defy representation by the traditional modeling paradigms: they combine many different scales, discrete and continuous behaviors, individual autonomous actors, aggregate feedback systems, and uncertainty.

CCIM Makes Significant Impacts in National Climate Consortium: The DOE Scidac Climate Consortium project is focused on developing a highly scalable climate model with a full carbon cycle. Scalability has motivated us to evaluate the use of cubed-sphere dycores in the atmospheric component (CAM), while the emphasis on carbon cycle modeling and the associated chemistry requires a dycore with a locally conservative, positive-definite advection algorithm. As research on positive-definite dycores progresses, we have been evaluating CAM using the HOMME-SE (spectral element) dycore. Initial runs of CAM with HOMME-SE showed that the dissipation mechanisms in HOMME-SE, were far too viscous when there was a strong enstrophy cascade typical of realistic climate simulations. To remedy this problem, we implemented a hyper-viscosity term that replaces the conventional viscosity and the element based filter used in HOMME (motivated by the fact that hyper-viscosity has been used in CAM since its inception as CCM0). This work required the development of a weak-form-divergence operator, which we also used in the continuity equation to achieve exact, local conservation of mass and exact conservation of total energy. Both of these improvements have been tested in the shallow water equations and we are currently evaluating their performance in CAM-HOMME using the aqua planet model problem. Our results using the hyper-viscosity method on the vortex-breaking problem from Polvani, et. al., JAS 1996 are in excellent agreement with his work.

XYCE Physics Additions: First ever, physics-based, transient neutron compact model for bi-polar transistors as part of the QASPR project. This model combines multiple 0-D reaction regions, all coupled with diffusive transport, to the industry standard Gummel-Poon compact model. This model provides Sandia with the ability to predict neutron damage effects at the transistor level as well as the efficiency necessary to simulate full circuit calculations as required for weapons component-design and qualification to hostile environments.

Charon Physics Additions: The Charon team has made tremendous progress in both implementing the full-physics defect package for QASPR calculations as well as becoming one of Sandia's preeminent high-performance computing applications. Charon is being routinely used to perform UQ calculations for the QASPR Prototype activities as well as large-scale computations on Red Storm and ASC Purple (LLNL) including runs on up-to 12,000 processors. Additionally and given these accomplishments, Charon has now become an official ASC application code for FY08, previously having been supported by CSRF and QASPR and research funding from LDRD.

XYCE to Impact US Industry: Xyce has now been licensed by an EDA startup company (Fastrack Design, San Jose, CA) and we expect the details of the agreement to not only help the Xyce technology impact US industry but also develop a collaboration with Fastrack to improve Xyce for Sandia and provide an increased level of production support to internal customers.

CCIM Supporting COIL Laser Development: The team working on the Chemical Oxygen-Iodine Laser (COIL) for the Air Force Research Laboratory (ARFL) has made substantial progress in modeling the Singlet Oxygen Generator component of the laser. Their initial successes have impressed the Air Force customers and resulted in increased funding for FY07 as well as a three-year funding stream initiated by the AFRL for continued support.

CCIM Organization: The following is an overview of who we are, what we do, and how we fit into the larger picture.

Sandia

Sandia is a mission-driven national security laboratory. We are managed by Sandia Corporation, a not-for-profit, wholly owned subsidiary of Lockheed Martin Corporation, for the National Nuclear Security Agency (NNSA), an independent agency within the U.S. Department of Energy (DOE). Along with the other DOE labs and the Jet Propulsion Laboratory, we are an example of a Government-Owned Contractor-Operated (GOCO) organization.

We have two main sites: the large majority of our technical efforts are in Albuquerque, NM, with over 8000 staff and contractors. We also have somewhat less than 2000 staff and contractors at our site in Livermore, CA. Sandia's missions include responsibility for all non-nuclear systems in the nation's nuclear weapons; related nuclear safety responsibilities; a large set of energy and critical infrastructure programs; work for the DoD; work with the U.S. national intelligence community; and work with U.S. industry. About half of Sandia's funding comes directly from NNSA, about 20% from energy programs, about 20% from the intelligence and defense communities, and about 10% from U.S. industry. The laboratory's overall budget exceeds two billion dollars.

A key new mission for Sandia is as a science and technology laboratory for The Department of Homeland Security. To date our principal efforts for DHS are focused on biological, chemical and nuclear terrorism, as well as sensors and sensor networks. However, DHS is facing organizing challenges and, to-date, has put little emphasis on R&D efforts. Currently, Sandia and LANL head NISAC, National Infrastructure Simulation and Analysis Center (NISAC), which has become part of DHS. A new building at Sandia (NM) for NISAC was dedicated on September 18, 2006.

Our Competitors and Cohorts

As Sandia is an applied science and technology laboratory, so is our center, the Center for Computation, Computers, Information, and Mathematics - an applied High-Performance Computing and Technical Informatics (HPCTI) R&D enterprise.

As a laboratory and as an HPCTI enterprise we see ourselves as a leading mission-driven R&D center. Our cohorts within the national R&D enterprise include LANL, LLNL, ORNL, LBNL, JPL, ANL, PNNL, INEEL, NIST, and MIT's Draper Lab.

Computing and Information at Sandia:

The Computational Science, Information and Technology (CSIT) enterprise at Sandia is divided as follows:

- The CIO functions (Centers 9300, 9500, 9600, half of 8900; ~550 staff):
    - Network and security infrastructure deployment;
    - Central computer operations;
    - Desktop services;
    - Desktop services for 25,000 desk/laptops;
    - Oracle and PeopleSoft development, deployment and operations;
    - Knowledge management.
- Software engineering and real-time systems development (Center 5500 and part of Center 8900; ~150 staff)
- Energy, Resources and Systems (Center 6300, ~100 staff computing staff)
- National Intelligence Computing and Signal processing (Center 5900, 25-50 computing staff)
- Engineering Simulation Development and Application (Center 1500, ~200 computing and analyst staff)
- R&D for high performance computing, computational science, and technical informatics –The Center for Computation, Computers, Information and Mathematics (CCIM, Center 1400, half of Center 8900; ~300 staff, limited-term employees, and full-time contractors)

This year, you are reviewing the R&D programs of the CIS Council, which comprises all of the work in Center 1400 (CCIM) and about half of Center 8900's efforts. CCIM comprises three groups as shown in the organization chart included in the CCIM overview presentation:

1410: Computer Science and Mathematics (David Womble)

1411 Optimization and Uncertainty Quantification (Jim Stewart)

1412 Computational Biology (Danny Rintoul)

1414 Computational Mathematics and Algorithms (Scott Collis)

1415 Discrete Algorithms and Math (Suzanne Rountree)

1416 Applied Computational Methods (Andrew Salinger, acting)

1420: Computer and Software Systems (Sudip Dosanjh,)

1421 Computational Modeling Sciences (Teddy Blacker)

1422 Scalable Computer Architectures (Jim Ang)

1423 Scalable System Software (Neil Pundit)

1424 Data Analysis & Visualization (David Rogers)

1430: Computer Science R&D (Jennifer Nelson)

1431 Computational Shock and Multiphysics (Randy Summers)

1433 Exploratory Simulation Technologies (Jim Strickland)

1435 Multiscale Dynamic Materials Modeling (John Aidun)

1437 Electrical and Microsystems Modeling (Scott Hutchinson:)

There have been several management changes in the past year in addition to changes already noted: David Rogers and Jim Stewart are the manager additions Scott Mitchell rotated back to staff.

In addition to these three groups,

David Womble is the Director of the Computational Sciences Research Institute (CSRI).

John Brewer runs the Center's business administration office (1054).

M. Danny Rintoul manages the New Initiatives effort within CCIM (Dept. 1409).

The Center has three Senior Scientists, Bob Benner, Jim Tomkins and Tim Trucano and 17 Distinguished Members of Technical Staff.

Our total annual funding is in the range of $50M –$60M for staff and equipment, not including platforms capital equipment.

CCIM History: The evolution of CCIM's Mission

Our mission has evolved as computing has matured at Sandia. When Ed Barsis founded this effort in the mid-eighties, Sandia was a second-tier computing lab. We looked to LANL and LLNL for leadership in this area. At that time we recognized several key factors that drove us to change our approach:

1. Our need for high-end computing was growing exponentially;
2. The need could not be met if we followed the cost curve of vector supercomputing;
3. Commodity parts were going to change the world;
4. Caltech and a few small companies were exploring massively parallel (MP) programming approaches, which I assessed would change the world;
5. Except for a few "nuts," people were ignoring MP Processor computers (MPPs);
6. There was a real opportunity for Sandia to rapidly establish leadership in HPC by changing the rules of the game.

We asked for and were given the initial mission of creating a whole new way of doing computing at Sandia that would put us on the path to provide the most powerful computing capabilities in the world for SNL's national security missions.

To that end we created interdisciplinary teams that systematically attacked the problems standing in the way of making MP programming succeed.

Initially we focused on two things:

1. Proving that MP programming worked on real scientific and engineering applications - believe it or not there was a huge amount of skepticism in the community, including within SNL. To that end, we set up teams who - over the period 1987 - 1992 – literally developed highly scalable MP program versions of just about every type of code Sandia was interested in. This amounted to perhaps 100-150 man-years and resulted in several dozen new MP codes that proved (to us anyway) the viability and usefulness of the MPP approach. It also fostered major advances in algorithms and libraries.

2. Evaluating various architectural choices: our first machines were a 1024-processor nCUBE10, arguably the first true MPP competitive with vector supercomputing, and a 512-Processor (16384) CM-2. This work convinced us that the future lay with MIMD and shared memory machines rather than SIMD machines.

Our second-generation machines were two 1024-PE nCUBE-2's. These machines became workhorses for certain groups within Sandia. They performed outstandingly. Their balances and reliability were phenomenal: at one point one of them went nearly a year without an unplanned interrupt! We also evaluated the BBN Butterfly, the IPSC-2 and IPSC-860, the Kendall Square, the AMD, among others. We also experimented with transputers and built two generations of multi-threaded H/W and S/W architectures, including building OSs and functional language compilers [This latter work led to successful collaborations with Motorola, Arvind and Los Alamos.] One thing we learned was that with then-current technologies Dataflow didn't provide enough benefit to justify its radical new approach.

This evaluation work was key to forming our approach to MPP computing:

1. Keep it simple—most fancy architectural features actually detracted from performance and reliability and codability.
2. Balance is critical to scalability
3. For a scalable system to succeed not only the code and the hardware must scale but also the systems S/W and the RAS features must also scale.

We succeeded better than we expected in this first stage. By 1993, we had carried out all the prototyping efforts we needed and felt that it was time to move MPP into prime time.

To that end, we partnered with Intel to field the first Paragon system: a 3700-PE system with a peak of over 100 GFLOPS. For the first time, in 1993 Sandia operated the world's fastest computer (measured by the Top 500 list.). In the intervening 10 years, Sandia has operated the largest and fastest machine for parts or all of 6 years (the Paragon for parts of two years and the Intel TFLOPS for the better part of 5 years). Red Storm now tops the competition in 4 out of 6 of the newly adopted measures of supercomputer performance:

We also began to develop the *enabling technologies* that are needed to make a full production environment:

- Representing complex 3-D geometries;
- Meshing those geometries;
- Distributing the mesh onto a large MPP (Static load balancing);
- Dynamically re-balancing the application load as it evolves during the run;
- Developing scalable input/output capabilities;
- Providing high-efficiency, highly scalable solvers and other algorithms;
- Encapsulating these into robust libraries;
- Combining parallel methods with object-oriented methods.
- Continue writing and updating scalable MP codes for SNL's user/analyst community.

When we took delivery on the Paragon, we soon realized that we had a disastrous problem on our hands. Not only was the OS supplied by Intel very large (it took over half the machine memory), but it also didn't scale—its image grew as the number of processors and its overhead did as well. To save ourselves, in about three months time we took work that we had been doing in a research effort on the nCUBEs - aimed at creating a portable highly-scalable, low-footprint "nano-kernel" operating system for MPP's - and ported it to the

Paragon as its compute node OS. This was the beginning of Sandia's initiatives in scalable OS and runtime systems. These efforts have formed the basis for the systems S/W environments on all subsequent Sandia Architectures: SUNMOS, PUMA, COUGAR, PORTALS (within Linux and within Cougar) and now CATAMOUNT. Versions of these OS's have worked effectively on the nCUBES, the Paragon, ASCI TFLOPS, and Cplant.

Our next big endeavor was to develop Cplant, which took Thomas Sterling's Beowulf ideas to the next logical step: we would make highly scalable, full-user-environment virtual supercomputers out of commodity cluster and semi-commodity interconnects. This work led to the current state of a 2.5 TF supercluster that provides about the same number of cycles to Sandia analysts as does ASCI TFLOPS – at about ½ the integrated cost of ownership. Our work has been put into the open source community and we have also provided it as the initial basis of a start-up company, Unlimited Scale Inc., headed by former Cray executives, John Rollwagen, Steve Oberlin and Mick Dunworth. The Cplant cluster effort has been retired now that commercial cluster computers have become well established, thanks, in part, to the accomplishments of the Cplant project. Our cluster efforts are now morphing to create an open-source LINUX compatible LWK based loosely on CATAMOUNT and to create a portal open-source RAS environment for clusters.

The late 90's: A Major Change in Direction:

By 1996, we determined that our simulation capabilities had matured to the point that in the simulation arena we were able to turn over most production code development to the Engineering Sciences Center (1500), along with many 1400-raised staff and managers to create an MPP-savvy culture in 1500. We began to concentrate on developing and prototyping *new classes* of simulations and on creating the *enabling technologies* that 1500's code developers and analysts, and 1800's materials, microelectronics and MEMS scientists would need. We also worked to empower the weapons component and system designers to carry out their early design efforts more efficiently by using design-appropriate simulation tools themselves early in the process rather than creating a design and asking analysts and/or testers to check it after it was a long way down the design/manufacturing road.

Beginning in about 1998, a number of us at Sandia realized that we were faced with an explosion of data, little ability to manage that data as information and less ability to turn that information into knowledge. Our mission as a lab was evolving: we were being pushed to help solve problems that increasingly involved creating, receiving, processing and managing huge amounts of data. We also realized that we were going to have to move into biology and biotechnology as a lab and that the initial grand challenges that we would have to face in biotech were informational rather than simulation. In the area of simulation, we realized that often the efforts devoted to technologies (especially interactive I/O, data viz, and meshing) were rate limiting rather than the execution time of the simulation itself.

Accordingly, in the period 1999-2005, under Bill Camp's leadership we did some new things:

1. We began to focus more heavily on information science and cognitive science.
2. We reinforced our commitment to leadership H/W and S/W architectures; we re-developed a competitive design capability for H/W architectures (Sandia has long been a leader in parallel embedded computing, e.g. for ATR and in signal processing computers, and had a competitive effort in dataflow architectures in the early 90's). We strengthened

our commitment to systems software. We felt that we had a unique role. We could work as a bridge between academic researchers who cannot build interestingly large instantiations of their idea and industry that cannot afford to take risks in architectural efforts. We also felt that our application knowledge put us in a good position to make architectural choices that would be helpful rather than hindering toward applications performance. We also wanted to do something about scalability and flexibility of I/O, and about reliability - two things that have suffered greatly in the commoditization and clusterization of HPC.

3. We began to play a much more proactive role within the external community - forging new partnerships with federal agencies (e.g. NSA), taking part in major community initiatives (e.g. Petaflops working groups), creating collaboration and partnering mechanisms for academia and industry (i.e. CSRI).
4. We designed a new generation MPP, Red Storm, that combines highly balanced, but largely commodity, architecture with inherent and explicit RAS features. Red Storm builds on the success of ASCI RED (a.k.a. ASCI TFLOPS) and the Cray T3E. It is designed to carry us architecturally to a Petaflops by 2010 or before. There now exist 25 Cray XT3/4 systems at 17 different sites.
5. Beyond Petaflops levels the MPP architecture will be under severe duress due to the large numbers of processors whose threads will have to be explicitly managed. Therefore we are actively pursuing partnerships in new "beyond-MPP" architectures. We are also building on our advanced software R&D capabilities and combining those with a re-invigorated hardware R&D capability. We will deploy these to develop and prototype architectural ideas.
6. In the applications arena we are moving further into cognition and perception sciences – as applied to national security. We are also developing more new predictive policy toolsets for economics, global climate systems and warfare. Our new emphasis on VLSI circuit and device-level simulation not only is providing a much-needed capability for Sandia's electronics designers and ultimately for US industry, but it also offers potential synergies with our architecture R&D.
7. Sandia is making a major push for achieving inertial confinement fusion (ICF) in the laboratory through our Z-pinch accelerator technologies. The principal analysis tools for that effort come directly out of CCIM in partnership with our fusion groups. They also provide the single largest projected application driver for our efforts.
8. In 2003 we formalized our diversification efforts by creating a New Initiatives Department. The manager leads and coordinates the Center's on-going efforts to bring HPC to bear on new technology areas.
9. The initial new initiative was biotechnology. This has given way to Climate Simulation, Agent-based methods, and informatics -- all of which present substantial demands on both hardware and software computing capabilities. They are, thus, strong drivers for advancing HPC capabilities. In turn, these areas stand to benefit greatly form successful application of HPC to their needs.