Advanced Computational Capabilities for Exploration in Heliophysical Science (ACCEHS) --- a Virtual Space Mission

A. Bhattacharjee,[1] K. Schrijver, W. Abbett, F. Cattaneo, L. Chacon, C. Cohen, W. Daughton, P. Erickson, L. Fisk, K. Germaschewski, J. Giacalone, T. Gombosi, M. Hesse, J. Huba, M. Hudson, V. Jordanova, P. Judge, H. Karimabadi, G. Lapenta, A. Lazarian, Y. Lin, J. Linker, H. Liu, J. Luhmann, M. Miesch, M. Moldwin, N. Omidi, T. Onsager, M. Opher, M. Oppenheim, J. Raeder, A. Ridley, P. Riley, R. Schunk, M. Velli, M. Wiltberger, and G. Zank

Executive Summary

In the dynamically complex, nonlinearly coupled domains of heliophysics, the efficient use of computers is as important as access to state-of-the-art *in situ* and remote-sensing instrumentation. Computers help us explore where we cannot sense, build a comprehensive view of very sparsely sampled environments, provide forward modeling where inversions fail, and put the heliophysical domain into a controllable laboratory setting. Computers are also are critical in processing the terabytes of data coming from both real-world and virtual-world experiments.

With the development of computer hardware supported by large economic interests in industry, we can leverage these investments with a fraction of their true cost to rapidly move beyond the decades-old legacy codes we now mostly work with, improve the foundations and capabilities of the underlying physical models, integrate across physical domains, and reach out across wide ranges in scales. This is societally relevant as we work towards more realistic forecasting of space and terrestrial weather and climate, capitalizing on advances in computer architecture and data storage to increase the capability of our workforce.

As discussed in a community-wide workshop that is the foundation for this concept paper, multiple critical problem areas in heliophysics are ready to be moved forward through advanced computational capabilities, with benefits across multiple sub-disciplines and with value to society. Examples of such transformational projects include:

- Generation and emergence of active regions from the convective envelope of the Sun;
- The evolution of the ambient solar corona and its coupling to the inner heliosphere;
- Understanding and parameterizing the kinetic solar wind from the inner to the outer edges of the heliosphere;
- Acceleration and transport of energetic particles in the heliosphere;
- Impact of severe storms on the geospace environment;
- Accurate models of the radiation belts;
- Ionosphere-thermosphere-magnetosphere (ITM) coupling;
- Prediction of communication outages due to ionospheric density irregularities and turbulence.

Recommendation: We recommend that NASA, perhaps in partnership with NSF and other agencies, should lead by establishing a new peer-reviewed program in which critical-mass groups of heliophysicists, computational scientists, and applied mathematicians can be brought together to address transformational science questions that support its flight missions and advance heliophysics. The level of synergistic collaboration of such groups, requiring $5-10M per project spread over several years, exceeds what is currently supported within GI, TR&T, HTP, or SR&T program envelopes. This new program should eventually support up to five groups with steady funding for periods of up to five years at an approximate level of $2M per year each. These groups should primarily pursue a scientific problem, advancing through reviewed milestones to design and use tools for theory, modeling, and data assimilation that will exploit the capabilities of state-of-the-art supercomputers. In doing so, they should also support the training of postdoctoral scientists and graduate students in advanced computational capabilities.

While the concept underlying this paper has taken on special urgency in recent years, the 2003 Decadal Survey Report emphasized the important role of computation in its "Coupling Complexity Research Initiative". That Report identified some key issues that remain as valid today as they were nearly a decade ago, including: (i) the importance of clearly articulated support and funding lines for model development, (ii) the need for computer hardware, which should be treated like hardware for experimentation, (iii) the necessity of science questions driving computational initiatives, and (iv) the importance of supporting data assimilation and exploration techniques in synergy with computational model development.

1. Background

Computational modeling, simulation, and data analysis have been among the most important drivers of scientific discovery during the last three decades. This progress has been enabled by remarkable leaps in computing technologies, producing parallel computers of great power and speed which have been brought to bear on increasingly sophisticated software and efficient algorithms. It is reasonable to anticipate a near-future transition from the present-day terascale and petascale systems to the exascale, which will empower the science community further to undertake challenges that can be potentially transformational.[2]

The heliophysics science community is experiencing a rapid and radical transformation because of vast increases in the sophistication of instruments and in data volumes. Progress in our understanding of heliophysical processes requires that data analysis be combined with computational modeling, numerical simulation, and data-assimilation programs. The effectiveness of our community-wide theory, modeling, and data assimilation and analysis efforts depends critically on the development of innovative numerical algorithms and their use on high-performance computing platforms, as we see happen in other scientific disciplines (such as astrophysics, high-energy and nuclear physics, plasma and fusion science, and climate prediction and change), which are well-positioned to exploit fully the power of new computing technologies. We cannot rely on the slow diffusion of the fruits of the efforts made by other scientific and engineering disciplines, but need to actively work on advancing our discipline's capabilities to meet the demands for scientific breakthroughs, the design of next-generation space missions, and to tap into the pool of new and young talent, eager to bring the power of new computing technologies and methodologies to bear upon heliophysical science challenges.

A community-wide Workshop on Advanced Computational Capabilities for Exploration in Heliophysical Science (ACCEHS), was held on August 16-18, 2010 at NCAR in Boulder, Colorado (http://www.hao.ucar.edu/ACCEHS/). The Workshop brought together over 80 scientists in heliophysical science, as well as experts from the climate and computer science communities. In what follows, we describe the key findings of the Workshop, cast in the form required by a concept paper for the Decadal Survey, recommending possible strategies for action summarized above. A more detailed Report for the ACCEHS Workshop is currently being written.

2. Key Challenges and Opportunities in Heliophysical Science Disciplines

The concept underlying this paper touches on all three elements of the present Decadal Survey: Atmosphere-Ionosphere-Magnetosphere Interactions, Solar Wind-Magnetosphere Interactions, and Solar and Heliospheric Physics. The primary goal of this section is to articulate challenges or questions where targeted investment in advanced computational capabilities has the potential to transform heliophysical science. Our goal is *not* to assemble a comprehensive list, and certainly not to imply a prioritization. The topics identified in this Section should be viewed as examples of challenges that can take heliophysical science to the next and higher level of discovery.

2.1 Generation and emergence of active regions from the convective envelope of the Sun

Magnetic activity on the Sun originates from hydro-magnetic dynamo action in its highly turbulent convective envelope. A unified, comprehensive understanding of the diverse range of magnetic activity exhibited by the Sun, from small-scale flux elements in the photosphere to global patterns of magnetic activity such as the sunspot cycle, remains an unfulfilled challenge. In particular, understanding how sub-surface magnetic flux emerges into the solar atmosphere, energizes the corona, shapes the heliosphere, and regulates space weather is an essential prerequisite in the effort to understand solar variability and its impact on the Earth and other planets of our solar system.

Modeling the generation and emergence of magnetic flux on the scale of active regions requires a realistic physical description of the complex boundary layers that straddle the Sun's convection zone. Near the base of the convection zone, large thermal, rotational, and magnetic gradients promote the generation, accumulation, and subsequent destabilization of magnetic structures that can buoyantly rise to the photosphere. As these structures pass through the visible surface and expand into the solar corona, they traverse a thin boundary layer where the physical environment changes drastically. Plasma densities drop by ten to twelve orders of magnitude, radiation and electron thermal conduction replace convection as the primary means of energy transport, and magnetism overtakes gas pressure as the dominant energy reservoir. The physics of this transition is particularly complex in the solar chromosphere, where shifting ionization states and non-local radiative processes become energetically important. Similarly, the physics of the boundary layer at the base of the convection zone is complicated by internal and interfacial wave

modes and tachocline instabilities.

The challenge, then, is to develop the theoretical and computational techniques necessary to model these physically, spatially, and temporally disparate regimes in a way that retains the essential physics well enough to interpret current and future ground and space-based observations. Currently, global convection simulations exhibit differential rotation and large-scale dynamo action but do not have sufficient spatial resolution to capture the formation and destabilization of concentrated flux structures. Local flux emergence simulations adequately capture the radiative MHD of the upper convection zone but rely on idealized initial conditions and simplified, spatially-limited models of the chromosphere and corona. A unified, higher-fidelity model is within reach but will require sophisticated, coupled, numerical algorithms that can fully exploit next-generation computational resources.

2.2 The evolution of the ambient solar corona and its coupling to the inner heliosphere

Many of the essential questions of solar physics are reflected in those about the structure of the solar corona. What mechanism(s) heat the corona? How is the fast solar wind accelerated? What is the origin of the slow solar wind? An understanding of coronal structure is not only important in its own right, but is implicit to understanding the geomagnetic effects of CMEs.

Phenomena in the solar corona and solar wind occur at many different scales. Observations of smaller scale phenomena give us clues to the fundamental processes that heat the solar corona and accelerate the solar wind. Models of the processes that drive these phenomena (e.g., reconnection, wave heating, and acceleration) are highly idealized but allow us to explore the viability of different mechanisms. Large-scale observations of the corona show us the consequences of the small-scale behavior –coronal streamers, coronal holes, the slow and fast wind. Global models of the corona and wind can capture much of this behavior, but must use empirical prescriptions or parameterizations of the smaller-scale physics to obtain realistic results – these constraints are imposed by the lack of resolution available to incorporate more basic physical ideas. These disjoint treatments make it difficult to test and improve theoretical models.

To understand which dominant physical mechanisms shape and power the solar corona and solar wind, global models must incorporate the small-scale processes as consistently as possible so that we can directly test theories with observations. Current state-of-the-art models cannot resolve the effects of random photospheric motions, which shuffle the footpoints of the coronal magnetic field and allow closed field lines in the streamer belt to interact with nearby open field lines in the coronal holes. Will significant reconnection occur? Can the resulting plasma release account for properties and extent of the slow solar wind? The introduction of these dynamics require ~10-100 times the number of grid points to resolve the subsequent scales. Breakthrough calculations (10^9 grid points and beyond) are becoming tractable as computers transition to megacore architectures. A major model development program is urgently required to take advantage of this opportunity, because present-day codes cannot be simply extended to the new state-of-the-art machines. Extensions to the physics of the models are also necessary to accomplish these groundbreaking simulations. For example, multi-fluid calculations with ionization equilibrium equations for the different species will allow the composition of the solar wind to be predicted from the models and tested against observations. Kinetic physics may very well be necessary to completely understand solar wind heating and acceleration, and may eventually need to be incorporated via coupling of kinetic simulations to MHD models.

2.3 Understanding and parameterizing the kinetic solar wind from the inner to the outer edges of the heliosphere

Although our well-developed and widely applied MHD 'fluid' concept of the solar wind has served us for decades as a reliable means of describing the macroscopic features of the interplanetary medium, it is inadequate to fully understand the details of the underlying behavior of the ions and electrons that make up the solar wind that the Earth is constantly exposed to. Numerical simulations of the kinetic processes involved in solar wind heating/acceleration and the wave-particle interactions that pervade the heliosphere are currently confined to local or highly simplified global treatments and cannot capture the effects of the boundary conditions and radial evolution that are so key to understanding the solar wind. Large-scale kinetic simulations of solar wind physics, accommodating more aspects of both ion and electron distribution functions in realistic coronal and interplanetary field geometries have the potential to revolutionize common working paradigms of heliophysics. For example, they can lead to improved parameterized descriptions of solar wind heating in the increasingly realistic global heliospheric simulations. They can lead to more accurate interpretation and use of ion and electron thermal anisotropies

and halo electron strahl.

At the farthest edges of the heliosphere, the solar wind interacts with the local interstellar medium where the distribution of both the plasma and neutrals can be highly anisotropic and non-Maxwellian, leading to significant effects on the global heliosphere. Moreover, from observations returned by Voyager and IBEX spacecraft mapping the heliosheath and the global structure, it is clear that the structure of the magnetic field plays a crucial role in this region. These observations reinforce the need to resolve disparate physical scales and kinetic processes. An example is the behavior of the heliospheric current sheet close to the heliopause. In the heliosheath, the sector regions approach each other as the solar wind slows down. The current state-of-the-art relies on idealized models that largely neglect the complexities of the heliospheric current sheet and the intrinsic time dependence of the solar wind. The demands of modeling extremely fine scales within a 3D time-dependent MHD code over very long times, couple with the need to model the kinetic physics of neutral hydrogen and plasma processes over several solar cycles makes this one of the most challenging computational problems in space physics.

2.4 Acceleration and transport of energetic particles in the heliosphere

Understanding how solar energetic particles (SEPs) are accelerated and transported in the heliosphere is a long and outstanding problem that is at the heart of space weather and cannot be attacked without a larger effort than is currently employed. Presently, only portions of the entire problem are being addressed through separate modest numerical simulations, e.g., energy release and magnetic reconnection at the Sun, shock formation, particle acceleration by interplanetary shocks and magnetic reconnection, and particle transport through the heliosphere. Coupling these simulations, e.g., incorporating particle acceleration into a simulation of coronal mass ejection eruption, has proven to be difficult and generally beyond the currently available resources (both in terms of computational resources and in terms of funding of a team of experts for a significant amount of time).

Similar to the problem of understanding the solar wind in the heliosphere, the acceleration of SEPs requires incorporating kinetic physics and a large range of physical scales into numerical simulations. A prime example is the question of how thermal, or very low-energy particles, that move slowly with respect to the shock are accelerated by the shock to high energies, is largely ignored in current modeling. This problem involves scales that range from the gyroradius of the thermal particles (~100 km at 1 AU) to the scattering mean-free path of the highest energy particles (~> 1 AU for ~GeV protons at 1 AU). The kinetic physics is important because the shocks that accelerate the particles move slowly enough that they are affected by the microstructure within the shock layer. Current models either ignore the shock microstructure and consider the acceleration to the highest energies only, or include the microstructure but are limited in the size of the model and therefore are limited in terms of the maximum energy achieved.

2.5 Impact of severe storms on the geospace environment

The magnetosphere represents the furthest extent of Earth's environment into the surrounding plasma and electromagnetic fields of interplanetary space. It stands between the solar wind and the ionosphere-atmosphere system and thereby controls the flow of mass, momentum, and energy. The magnetosphere spans a huge volume that is highly under-sampled, due to the small number of satellites making in-situ measurements and limited remote sensing opportunities. Understanding the magnetosphere is further complicated by the dominance of coupled, nonlinear physical processes that cover orders of magnitude in temporal and spatial scales. Conditions in the magnetosphere can change in a matter of minutes from quiet-time to storm-time, and storms can last from hours to days. Large geomagnetic storms can cause serious damage to technological systems like satellites, increase the radiation exposure of astronauts, and disrupt communication and power systems. Predicting geomagnetic storms (especially severe storms) is one of the biggest challenges in magnetospheric physics.

Given the complex and multi-scale nature of the magnetosphere and its coupling to the solar wind and the ionosphere-atmosphere system, understanding its dynamics requires state-of-the-art models, computational facilities, and access to comprehensive data sets. The two most promising approaches to understanding the global behavior of the magnetosphere are MHD/multi-fluid and hybrid simulations. MHD simulations have been in use for more than 3 decades and advances in hardware and software technology have resulted in their ongoing improvement. These models, however cannot accurately describe the inner magnetosphere because they do not include the energy-dependent drifts of the pressure bearing ring current particles. Robust coupling of MHD models with kinetic ring current models is difficult to achieve at present due to the required higher temporal and spatial resolution. Hybrid models treat the ions

kinetically and allow coupling of microscopic and macroscopic processes, but due to their considerably larger computational expense, hybrid simulations currently utilize a system size that is smaller than the magnetosphere. Both MHD/multifluid and hybrid models need to include ionospheric ion outflow models that will allow coupling between the magnetosphere and the ionosphere under different solar wind driving conditions. Advances in hardware technology and supporting computational libraries will allow future development of near-real-time predictive capabilities for geomagnetic storms.

2.6 Accurate models of the radiation belts

The dynamic variability of radiation belt electrons over orders of magnitude poses a challenge to modelers and a threat to our increased dependence on satellite systems vulnerable to these changes. A thorough understanding of the mechanisms involved in their access, transport, trapping, acceleration, and loss can be achieved only with a system level approach, including coupling between adjacent regions. Trapped MeV electron flux is highly variable, peaking during the declining phase from solar maximum, with strong enhancements around the ~ 11 year maximum in sunspot number. The low energy (eV) plasmasphere population which co-rotates with the Earth is continuously refilled by the topside ionosphere and stripped away by changing magnetospheric convection, typically around L ~ 4. Our understanding of the coupling of the solar wind drivers to the dynamics of the low energy plasmasphere population, and its effects on MeV electron fluxes, requires advanced modeling of wave-particle interactions near the plasmapause, responsible for determining the depth of penetration of MeV electron fluxes into the altitude range of, for example, GPS spacecraft. Recent observations of coherent large amplitude whistler waves in this region responsible for electron loss and local acceleration have been limited by time resolution of earlier measurements. Advances in spacecraft data acquisition at higher time resolution present a data storage and processing challenge similar to that faced by modelers.

Local micro-scale processes such as the generation of plasma waves and their effect on particle dynamics must be included in global models. Major computational challenges are thus to develop models that couple self-consistently the plasma and the fields across various regions of the magnetosphere, and that include both large-scale and micro-scale physics. This could be achieved by coupling global MHD and/or multi-fluid models with kinetic or full particle models, which each address different regions or physical processes. Including ionospheric outflow as well as solar wind entry in such models is essential, as such outflow/inflow repopulates the plasmasheet, the source of both energetic electron and ion components. Inclusion of plasmasphere dynamics in global models is essential, as the plasmapause has been shown to be a significant boundary affecting both radial transport and localized heating and loss of radiation belt electrons and ring current ions. Thus, particle energies span approximately six orders of magnitude over which particle dynamics must be modeled in a combined fluid and particle approach. Cross-coupling of particle populations over such a broad energy range, combined with the physical scale of the coupled system and range of timescales between global processes and localized acceleration and pitch angle scattering present major computational challenges. The development of data assimilation models will allow the integration of ground-based and space-borne data sets with models, soon to be augmented by the first spacecraft mission, RBSP, dedicated to radiation belt studies in two solar cycles.

2.7 Ionosphere-thermosphere-magnetosphere (ITM) coupling

It has become increasingly clear over the past several years that coupling processes between the lower atmosphere, thermosphere, ionosphere, and magnetosphere have a much larger impact on microscale to mesoscale behavior than previously understood. The major computational challenges are to develop accurate and robust numerical models for each region, and to seamlessly couple these models into a single, self-consistent unified model. The development of whole atmosphere models in recent years reflects these needs. Other examples include frameworks that link together models of the magnetosphere, ionospheric electrodynamics, and upper atmosphere.

An example of neutral atmospheric dynamics on the behavior of the ionosphere is the phenomenon of sudden stratospheric warming. During these events, lower atmospheric waves, such as planetary, tidal, and gravity waves, undergo significant variation and produce changes in the stratosphere, mesosphere, and thermosphere that significantly impact the electrodynamics of the ionosphere. Current modeling capabilities are inadequate to fully explain and predict these behaviors, in many cases due to the need for expanded resolution and improved boundary dynamics. Increased resolution and improved physics is critical to correctly quantify large scale waves (i.e., tides and planetary waves) and to resolve mesoscale

waves (e.g., gravity waves). These waves have important implications for ionospheric variability and space weather, but not reproducible by current global models.

2.8 Predict navigation and communication outages due to ionospheric density irregularities and turbulence

A classic example of this challenge is equatorial spread F (ESF), during which the equatorial ionosphere becomes Rayleigh-Taylor unstable: large (tens of km) electron density bubbles develop and rise to high altitudes (1000 km or greater at times). Attendant with these large-scale bubbles is a spectrum of density irregularities that can extend to wavelengths as short as 10 cm. Understanding and modeling ESF is important because of its impact on space weather: the associated electron density irregularities can cause radio wave scintillation that degrades communication and navigation systems.

The first major computational problem is to seamlessly couple different spatially overlapping, physics models. Large-scale (hundreds of m) ionospheric processes are well described by fluid theory. However, a kinetic description of the plasma is needed to model small-scale wave structures (~100 m) and the turbulent cascade of energy down to tens of cm. The computational algorithms to allow self-consistent coupling of fluid and kinetic models, which need to be developed, are crucial to describe the onset and decay of scintillation-causing irregularities. The second major computational problem is modeling the electrodynamics of the IT system. All global models of the IT system assume that the geomagnetic field lines are equipotentials; this reduces the potential equation to two dimensions. To accurately describe the self-consistent coupling of large scale to small- scale density irregularities, it is necessary to have a fully electrodynamics model. This requires the development of a fully three-dimensional solver that is robust, efficient, and uses parallel processing on a non-uniform grid. The third major computational problem is to couple computational models over a large range of spatial and temporal scales. The nominal spatial grid scales for current IT system models are 2.5-5.0 degrees in latitude and longitude (i.e., 300-500 km) and 2-10s km in altitude. The nominal time scale ranges from ~10s to ~10 min. To capture scintillation-producing ionospheric irregularities in a computational model, the grid resolution has to be reduced to ~1 m, which is about three orders of magnitude smaller than is currently feasible.

3. Computational Needs and Opportunities

Progress on some of the most exciting scientific challenges in heliophysics is awaiting the development of new computational techniques and the full utilization of petascale computing. Looking across the various sub-fields, the basic requirements for new tools and approaches share many common themes. Researchers seek to understand systems with vast disparities in spatial and temporal scales. This requires accurate models governing the large-scale evolution, which may depend on a kinetic processes. New petascale machines and other advanced hardware such as general-purpose graphical processing units (GPU) offer the potential for tremendous progress on these multi-scale problems. The ability to fully exploit these computers will permit calculations (100-1000) times larger than previous efforts. However, the spatial and temporal scale separations are so large that simply adapting existing algorithms to the new hardware is not sufficient. New breakthroughs will also require improvements in multi-scale algorithms and adapting the modern advancements in algorithms. In addition, new approaches are needed to deal with the large quantity of data generated by both simulations and observations. We have identified four basic categories of computational tools to address these needs:

3.1 High performance solvers and advanced algorithms for fluid simulations

Fluid simulations play a critical role in all sub-fields of heliophysics, including neutral fluids in the atmosphere/thermosphere, electrostatic fluid models in the ionosphere, and MHD models in magnetospheric, interplanetary and solar physics. These range from 1D regional models (such as solar wind acceleration), to 2D codes (such as magnetotail reconnection models) to fully 3D global codes (such as global magnetosphere and heliosphere codes). Most of these codes were written by scientists with deep understanding of the physics but limited background in modern algorithms, and before massively parallel machines were widely available.

Efforts to develop modern fluid simulation capabilities should include three key components: improved physics and closures, advanced solvers and the ability to fully utilize existing and future computer hardware. Depending on the specific science focus, the optimum mixture may differ. For example, how to properly capture kinetic effects within fluid calculations or describe the effects of turbulence remain key issues for accurately modeling magnetic reconnection, but improvements in solver technology and/or petascale computers are also very important for performing the high-resolution

simulations required to test improved fluid closures and subgrid methods. Most heliophysics simulations are physically far under-resolved, and would benefit greatly modern adaptive mesh refinement (AMR) methods and implicit time stepping. However, in order to simulate physically relevant large systems, the parallel and algorithm scalability is crucial. For explicit codes, parallel scalability is typically quite good using simple domain decomposition, and remains good for enhanced methods like block-structured AMR. While more difficult, recent work has demonstrated that implicit MHD algorithms showing such optimal algorithmic properties are possible. These issues are also important for electrostatic fluid treatments in the ionosphere, where there is a critical need for highly scalable 3D Poisson solvers on non-uniform meshes.

3.2 Advanced kinetic simulation approaches and algorithms

While fluid models have been very successful in modeling the macroscopic structure of many phenomena, there are many important science questions in heliophysics that require a kinetic description. A range of different kinetic simulation approaches have been employed including fully kinetic (ions and electrons), hybrid (fully kinetic ions and fluid electrons) as well as gyrokinetic models. Most kinetic simulations still employ a uniform mesh, with simple explicit algorithms. On the positive side, these simple well-tested algorithms can achieve excellent performance and scaling on modern computer architectures. Recent 3D fully kinetic simulation studies of magnetic reconnection have employed over ~1 trillion computational particles running on 10^5 computational cores, with an additional factor of ~10 anticipated in the next few years. These capabilities are leading to a range of new insights into the physics of reconnection, and may be extremely useful for addressing basic physics issues in shocks and turbulence. However, for many problems the spatial and temporal scale separations are so large, that simple explicit approaches (with uniform mesh) will always be limited due to rigid stability constraints and the long-time scale accuracy of the solution. While investments are clearly needed to help researchers exploit new architectures, there is an equal need for targeted investments in algorithms and multi-scale techniques. This includes, asymptotic-preserving discrete numerical formulations, non-uniform structured mesh, AMR, and implicit time stepping and/or discrete-event multi-stepping techniques. The scientific payoff will be high if researchers can effectively combine the raw power of petascale computing with modern advances in algorithms.

3.3 Computational framework: coupling physics across disparate scales

Development of a general simulation approach to handling multi-physics, multi-scale problems described by different models (codes) remains an outstanding research topic in computational physics. Problems in heliophysics will likely require software frameworks that include a superstructure layer that drives the coupled-model application, and an infrastructure layer that provides utilities and data structures for model developers. There are an abundance of heliophysics applications that will benefit from new investments in these capabilities. As one example, researchers have already successfully coupled ionosphere/thermosphere dynamics to global magnetospheric MHD codes. With petascale computers, global hybrid calculations are becoming increasingly realistic, but will still need to be coupled to ionospheric models. While hybrid calculation can properly treat the structure of ion-scale transition layers and shocks, neither approach (hybrid or MHD) can correctly capture the microphysics of magnetic reconnection. Future efforts in computational frameworks may allow researchers to couple more complete kinetic descriptions within the thin current sheets, and thus move towards a more predictive understanding of the global dynamics.

3.4 Analysis tools, parallel visualization and data mining and assimilation

Both simulations and space observations are producing increasingly larger data sets that are challenging for researchers to study with existing tools and approaches. New missions such as the Solar Dynamic Observatory have a data rate of 1.4 terabytes per day, while recent 3D kinetic simulations using the newest petascale computers are generating ~40 terabytes in a single run. Extracting new science from these enormous data sets is becoming progressively more difficult. Researchers are drowning in the over abundance of data, and legacy tools and approaches have failed to keep pace. The largest simulations are preformed at a few leadership class facilities around the country, and the resulting data sets are now far too large to move. To utilize these new machines, researchers need parallel tools to remotely visualize and manipulate the data, and to easily allow access to teams of researchers working in a collaborative fashion.

Although there have been efforts in recent years in extending simulation capabilities to petascale computers, less attention has been paid to the data analysis and visualization of the resulting large data sets.

In terms of spacecraft data, researchers have introduced data mining techniques and have worked with experimentalists in their analysis. This effort has met with considerable success and is rapidly gaining adoption. However, funding is required for further expansion of these techniques, as well as the integration of these new methods with the existing data analysis software used by experimentalists (e.g., autoplot, VO, etc.).

References

1. *Contact information*: Space Science Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, e-mail: amitava.bhattacharjee@unh.edu.
2. See the "Scientific Grand Challenges Workshop Series" on extreme scale computational science, organized by the Office of Science, Department of Energy (http://extremecomputing.labworks.org/index.stm).