MRI: Acquisition of a High-Performance Parallel Computing Cluster for Astrophysics Research at UC Santa Cruz

A proposal to
NSF Program Solicitation 11-503, Major Research Instrumentation,
Division of Astronomical Sciences

from the
Departments of Applied Mathematics and Statistics, Astronomy and Astrophysics, Computer Science, Earth and Planetary Sciences, and Physics at the University of California, Santa Cruz

Principal Investigator: Piero Madau
Co-Investigators: E. Asphaug, N. Brummell, J. Primack, S. Woosley
Senior Personnel: S. Brandt, C. Conroy, J. Fortney, P. Garaud, G. Glatzmaier, M. Krumholz
G. Laughlin, D. Lin, E. Miller, F. Nimmo, E. Ramirez-Ruiz

January 26, 2012

Contents

1 Project Summary **3**
- 1.1 Intellectual merit 3
- 1.2 Broader impact 3

2 Project Description **4**
- 2.1 Instrument Location and Type 4
- 2.2 Research Activities to be Enabled 4
  - 2.2.1 Background 4
  - 2.2.2 Results from prior NSF Support - NSF MRI grant AST-0521566 5
  - 2.2.3 Proposed New Research in Astrophysics 10
  - 2.2.4 I/O and Storage Management 14
- 2.3 Impact on Research and Training Infrastructure 14
- 2.4 Description of the Research Instrumentation and Needs 16
  - 2.4.1 Instrumentation 16
  - 2.4.2 Cluster Interconnects 17
  - 2.4.3 Compute Nodes 17
- 2.5 Management Plan 18
  - 2.5.1 The Site 18
  - 2.5.2 Machine Management and Time Allocation 18

3 Data Management **19**

4 References Cited **21**
4.1 Publications Enabled by Pleiades since 2007 21
4.2 Other References Cited 30

5 Budget Justification **32**

6 Facilities, Equipment, and Other Resources **33**

1 Project Summary

1.1 Intellectual merit

State of the art computational resources have been instrumental in making the University of California at Santa Cruz (UCSC) one of the world's leading centers for research in numerical astrophysics and planetary science. The machine proposed here will provide an order of magnitude improvement in our ability to address some of the most fundamental scientific questions of our time, from the structure of the early Universe and the nature of dark matter, through the assembly of the Milky Way, to the formation of extrasolar planets and our own solar system. The proposal also provides a new, substantial storage system that will not only support the ongoing research on the new machine, but also serve as an archive for datasets produced on larger off-site computers. The principal users will be the five co-PIs (Madau, Asphaug, Brummell, Primack, and Woosley), 11 senior associates (Brandt, Conroy, Fortney, Garaud, Glatzmaier, Krumholz, Laughlin, Lin, Miller, Nimmo, Ramirez-Ruiz) and over 80 graduate students and postdocs. This diverse group of researchers in astrophysics and planetary science spans the Departments of Applied Mathematics and Statistics, Astronomy and Astrophysics, Computer Sciences, Earth and Planetary Sciences, and Physics. The hardware proposed is a 30 Tflop machine with 200 conventional compute nodes (2400 general-purpose processing cores), 40 Graphic Processing Unit (GPU) accelerated nodes with 12 cores and 1 GPU unit in each, and 1 petabyte of fast storage. A variety of highly parallelized codes will be ported to this system to simulate phenomena such as cosmological structure formation, galaxy evolution, exploding stars, black holes, turbulence and magnetic fields, star and planet formation, and asteroid impacts. The new codes will be optimized by computer scientists in order to improve the efficiency of scalable storage systems.

1.2 Broader impact

The new cluster will enable cutting-edge science, help develop the next generation of simulation codes for astrophysics, provide a stepping stone to national leadership-class facilities, and promote the training of the next generation of computational astrophysicists. Our hybrid design, which couples the computer to a scalable storage database and an ultra-high-speed fiber-optic connection will also allow access to and analysis of petabytes of legacy data by the community at large. The combination will further cement UCSC's position as a center for computational astrophysics and a portal for simulation results and educational resources. It will facilitate the recruitment of excellent faculty, postdocs, and graduate students in astrophysics and planetary science, increase their collective productivity, and foster interdisciplinary collaborations. Studies of storage access patterns and efficiency by the computer scientists will be widely applicable to other storage systems outside of astrophysics. Indeed, one of our goals is the nurturing of a new kind of interdisciplinary computational scientist, well versed in both astrophysics and the computational tools of large-scale simulation and analysis. Team members will conduct an aggressive knowledge transfer program, providing the broader public with the opportunity to do and see astrophysics. The visual appeal of astrophysics simulations will be used to engage students in research, connecting communities as disparate as astrophysics and the visual arts. We will also disseminate the scientific images and insights of the Center via partnerships with high schools, planetaria, the NASA Ames Hyperwall, GalaxyZoo, and GoogleSky. Participants in astrophysics summer schools at UCSC will be given access to the cluster and therefore obtain unique hands-on experience with parallel supercomputing. This has been very successful in the past, for example, for the International Summer Institute for Modeling in Astrophysics (ISIMA) and the High-Performance AstroComputing Center (HiPACC).

2 Project Description

Title: Acquisition of Beowulf Cluster for a Center for Computational Astrophysics at the University of California, Santa Cruz

2.1 Instrument Location and Type

Location: University of California Santa Cruz
Code: MRI-31

2.2 Research Activities to be Enabled

2.2.1 Background

UCSC is one of the world's leading centers for research in computational astrophysics. At any time, about twenty faculty in four departments (Applied Math & Statistics, Astronomy & Astrophysics, Earth & Planetary Sciences, and Physics) and at least fifty graduate students and postdocs are using a variety of on-campus and off-campus computer facilities to address some of the more fundamental scientific questions of the time, from the structure of the early Universe and the nature of dark matter, through the formation of galaxies, stars, and planetary systems, to impacts in the solar system. Over longer periods, many more graduate students and postdocs are involved. Our papers are highly cited and our simulations are frequently featured in the media.

The primary on-campus facility currently used by this research cadre is *Pleiades*, a machine purchased through a previous NSF-MRI grant (see §1.2.2 below), that is approaching the end of its useful life. The machine has served our community well for about six years, but warranties have expired, the hardware is starting to fail and the technology is now outdated. The machine should be replaced; a machine 5 to 10 times more powerful can sit in the same footprint of space, power and cooling.

Pleiades has served a broad spectrum of needs including facilitating code development and providing a platform for small- and medium-scale experimental simulations in 2D and 3D, data analysis, and student training. It has allowed us to hone our tools for high resolution production 3D simulations that use tens of millions of CPU hours on much larger systems off campus. As a direct consequence of the demonstrable efficiency of these codes, our faculty have been successful in competing for time on some of the world's most powerful machines, specifically those owned by the DOE, NSF, and NASA. Unfortunately, these leadership-class facilities are not intended for, and actually discriminate against the midrange user. They are thus not conducive to student training, small scale exploratory calculations, or the dissemination of simulation data to the larger community. Much, maybe even the majority of astrophysical simulation with high scientific impact is still being carried out on mid-range clusters like the one we propose to purchase here.

The nature of of how we interact with computers is changing though. Increasingly, productivity is limited by the ability to analyze the large sets of complex data generated by our simulations. Our largest numerical simulations generate many tens (sometimes hundreds) of Tbyte of data. Data analysis must often be performed while running the simulations themselves, since the output data are too large to be moved or stored for reuse. Simulation data at remote supercomputer centers have a finite lifespan. As a result it is often difficult, if not impossible, to confront model predictions with the observational data. The issues of storage and archival of such large quantities of data and efficient access to such enormous datasets for analysis are becoming as critical as the calculation of the data itself. New scalable, heterogeneous, hierarchical methodologies appropriate for the scientific context are needed.

The new mid-size computing cluster we are proposing will have 6 times the computing power and 20 times the storage capability of the old *Pleiades*, while fitting comfortably within the same footprint of power, cooling and administrative support. It will once again enable cutting-edge astroplanetary science, allow the development of the next generation of codes, provide a steppingstone to national leadership-class facilities, and promote the training of the next generation of computational astrophysicists. In particular, this time, between 10 and 20 percent of the compute nodes will be GPU accelerated, to allow preparation for the forthcoming petaflop-scale GPU supercomputers at national centers like *Blue Waters* at NCSA and *Titan* at ORNL. The new machine's hybrid system design will integrate a high performance computing (HPC) cluster with a scalable storage database and the UCSC/CENIC ultra-high-speed Internet fiber-optic connection. This will allow access to and analysis of petabytes of legacy simulation data by the astrophysics community at large. In tandem, data access patterns will be instrumented to provide insight for research in scalable storage systems.

2.2.2 Results from prior NSF Support - NSF MRI grant AST-0521566

Title: Acquisition of Beowulf Cluster for a Center for Computational Astrophysics at the University of California, Santa Cruz
PI: Stanford E. Woosley
Co-PI: Pascale Garaud, Donald T. Gavel, Gary A. Glatzmaier, Joel R. Primack
Other Senior Personnel: Erik Asphaug, Douglas N.C. Lin, Piero Madau, Claire E. Max, Jerry E. Nelson
Amount: $1,100,000
Duration: 9/15/2005 – 8/31/2007

This previous MRI grant was used to purchase *Pleiades*, a Beowulf cluster with 828 cores, 1.66 TB of memory, a peak speed of 5.9 Tflops, and 55 Tbytes of storage. As of today, a total of 138 faculty, postdocs, and graduate students have an account on *Pleiades* and have kept it running at nearly full capability since its inception. Altogether, more than 200 papers in computational astrophysics and planetary sciences have been published thanks to *Pleiades* (listed in References Cited). Here, we provide some highlights of the research enabled by the previous grant, and emphasize the broad community that it has served.

Galaxies and Cosmology Research in galaxy formation was carried out by Madau and Primack and their students and postdocs. Madau's group included graduate students B. Anderson, X. Chen, J. Guedes, M. Kuhlen, V. Rashkov, and A. Sesana, and postdocs J. Diemand, Q. Yu, and M. Zemp. Research focused on the *Via Lactea Project*, a suite of the most extensive cosmological simulations ever carried out of the assembly of the cold dark matter halo of our own Milky Way (26; 27; 28; 101; 103; 179). Although this massive computational effort used large allocations of resources on NASA and DOE-provided high-end supercomputing, access to *Pleiades* for code testing and data analysis was essential. The project became one of only ten computational projects nationwide featuring in the *OASCR Breakthrough 2008 Report on Recent Significant Advancements in Computational Science*. The simulations show the existence of thousands of very concentrated, possibly annihilating (2; 102; 105), dark matter clumps surviving within the inner 50 kpc from the Galactic Center. *Via Lactea II* simulation data have been used to probe reionization and the "missing Milky Way satellite problem" (127; 128; 133; 155), to determine the velocity distribution function of dark matter particles for prediction of the dark matter–nucleon scattering rates in direct detection experiments (106), and to study the kinematics of recoiling massive black holes in galaxy halos (61). In the spirit of a "legacy" simulation, all processed *Via Lactea* data have been made available to the public (see http://www.ucolick.org/∼diemand/vl).

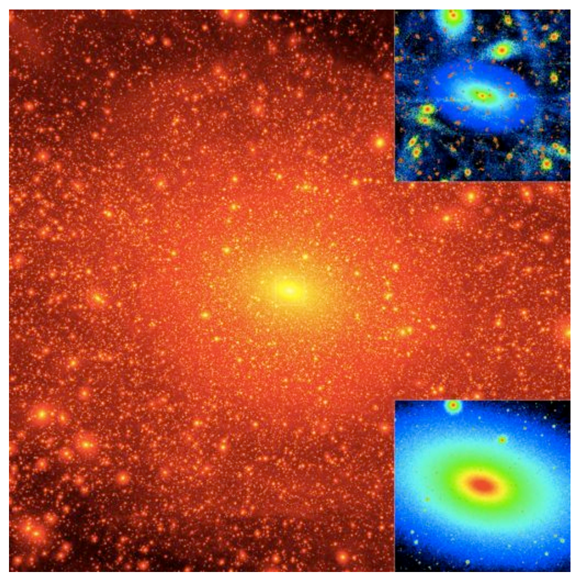

Figure 1: The *Via Lactea II* simulation of the dark matter halo of a Milky Way-sized galaxy. Projected dark matter density-square at the present epoch is shown in a 800 kpc cube. The insets focus on an inner 40 kpc cube, in local density (bottom), and in local phase space density (top). The *Via Lactea II* simulation has a mass resolution of 4,100 $M_\odot$ and a force resolution of 40 pc.

Primack's group included postdocs P. Jonsson, J. Lotz, and A. Romanowsky, graduate students B. Allgood, M. Covington, T. J. Cox, A. Dominguez, M. Fumagalli, R. Gilmore, G. Novak, and C. Pierce, and undergraduates A. Breslin, M. Coleman, T. Davalos, S. Kahn, C. Mazakas, M. Planta, and M. Rocha. A major program of research involved hydrodynamic simulations of galaxy mergers (21), including the effects of stellar evolution using the SUNRISE code (75), to determine the observational timescales of various morphological signatures (123; 124; 125; 122). These simulations were also used for predictions of galaxy lensing observations (139), for comparison to kinematic observations of galaxies out to redshift $z \sim 1$ (19), to develop and normalize an analytic model of galaxy mergers (18; 20). The *Bolshoi* large cosmological simulation suite (87; 151; 188; 13; 14; 156), while run on NASA supercomputers, has immensely benefited from code development and data analysis on *Pleiades*. Calculations of cluster shapes have been compared with X-ray observations in (38). A lot of effort was also devoted to semi-analytic modeling of the evolving galaxy population as a whole (177) and of the resulting extragalactic background light for comparison with observations (32; 54; 51; 52; 53).

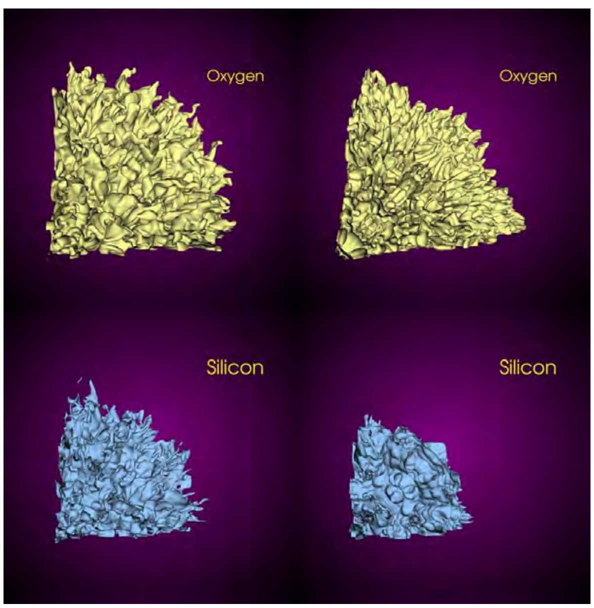

Figure 2: Mixing in a Type IIp supernova calculated in 3D using the CASTRO code (C. Joggerst, 2011 PhD thesis). The figure shows the Rayleigh-Taylor instability that develops as the helium core plows into the hydrogen envelope. Calculations were done at LANL following extensive developmental work on *Pleiades* (71; 72; 73).

High Energy Astrophysics Woosley's work focused on supernovae of all kinds and involved both code development and model building. In addition to Woosley himself, five postdocs (Yoon, Kasen, Ma, Malone, and Moll), four graduate students (Joggerst, Ma, Roberts, and Lovegrove), and one staff member (Dong) used *Pleiades* to study a wide variety of supernova science. Mixing and fall back in Type II supernovae was the topic of C. Joggerst's 2010 PhD thesis (Fig. 2; 71; 72; 73). Hubble Fellow D. Kasen used *Pleiades* to further develop the SEDONA code for computing supernova light curves and spectra, and applied it to the study of Type Ia supernovae (82; 83), core-collapse supernovae (81) and pair-instability supernovae (84). S. C. Yoon used Glatzmaier's 3D MHD code to study magnetic torques in differentially rotating massive stars (197). L. Roberts used *Pleiades* to develop neutrino physics modules for his protoneutron star evolution code (161). Worthy of special note was the developmental work on the radiation-hydrodynamics CASTRO code done by H. Ma and Woosley on *Pleiades* and their simulations of Type Ia supernova explosions (Fig. 3). CASTRO was developed as part of a DOE SciDAC-sponsored Computational Astrophysics Consortium headed by Woosley. While the main hydrodynamics drivers were developed by applied mathematicians at LBNL (212), the modules for nuclear burning and flame propagation in CASTRO were developed at UCSC on *Pleiades* and tested there on 2D models. This year alone, Woosley's group has been awarded 34 M CPU h at NERSC on *Franklin* and *Hopper* and 42 M CPU h at ORNL on *Jaguar*, and will be an early user of the *Blue Waters* Cray in Illinois. The ease of access and rapid turnaround at *Pleiades* were essential to getting Ma's thesis done in a timely way and demonstrating that the codes were ready for the leadership-class machines.

Ramirez-Ruiz and his group - which includes three postdocs (De Colle, Porter and Kasen), and five graduate students (Guillochon, MacLeod, Navarrete, Naiman, Rosen) - have used *Pleiades* to construct numerical models of a variety of explosive phenomena including gamma-ray bursts (112; 152; 153), white dwarf mergers (25; 66; 168), tidal disruption flares (65; 154; 169; 170; 171) and neutron star mergers (86; 202; 160).

High energy astrophysics research using *Pleiades* also included searching for gamma ray pulsars using the observations of the *Fermi Gamma Ray Space Telescope*. This resulted in a *Science* cover paper (1) (see Fig. 4). A Constrained Local UniversE Simulation (CLUES) was also used by (22) to model gamma rays from dark matter decay and annihilation in the local universe. Madau, with graduate students Chen and Sesana, ran numerical scattering experiments to study three-body interactions between ambient stars and a massive black hole binary (16; 17). Madau and postdoc Yu followed the kinematics of hypervelocity stars ejected by the massive black hole at the Galactic Centre as a probe of dark halo triaxiality (199).

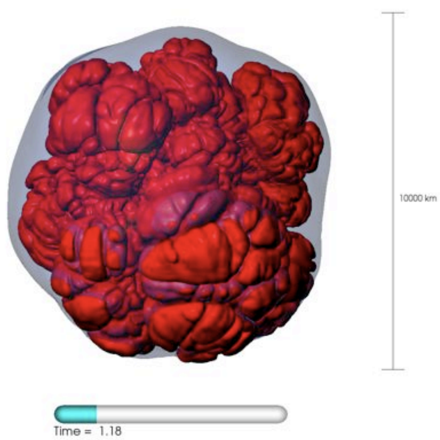

Figure 3: Centrally ignited carbon deflagration model for a Type Ia supernova (H. Ma, 2011 PhD thesis). 3D calculations were done using the CASTRO code at ORNL and NERSC following extensive developmental work on *Pleiades*. The figure shows the explosion of a Chandrasekhar mass carbon-oxygen white dwarf near the end of burning, 1.21 s after ignition. Red shows the interface between ash and fuel and light blue color indicates the star surface where density equals 10^6 g cm^{-3}.

Stars and the Sun Research in solar and stellar computational astrophysics spans three departments and broadly studies dynamos, mixing in stars and star formation. Brummell has been working on *Pleiades* with graduate student B. Byington on simulations of a novel "essentially nonlinear" and "stoked" dynamo mechanisms for magnetic field generation using Brummell's compressible MHD code called HPS. Brummell and Glatzmaier, together with graduate student K. White and postdoc C. Guervilly, have studied boundary-driven spherical dynamos related to the understanding of physical dynamo experiments, using both Guervilly's Boussinesq PARODY code and Glatzmaier's spectral code on *Pleiades*. Garaud and Brummell, with graduate student L. Acevedo-Arreguin and postdocs Guervilly and T. Wood, have also utilised a combination of Garaud's steady state solver, HPS and PARODY, extensively on *Pleiades* to formulate a new theory for the dynamics of the solar tachocline. Garaud, and former graduate student A. Traxler, postdocs S. Stellmach and T. Wood, and external students E. Rosenblum and G. Mirouh, have used a specially-developed code (PADDI) on *Pleiades* to study the transport properties of double-diffusive instabilities, leading to important discoveries about spontaneous layer formation in 3D and therefore mixing in oceanographic, planetary and astrophysical systems (46; 43; 181; 186; 44). Rogers and Glatzmaier have used *Pleiades* to help develop a new 3D global solar dynamo code that employs a spherical harmonics spectral method in latitude and longitude but a more flexible finite-difference method in radius, and solves an energy equation for temperature rather than entropy.

Figure 4: The search for gamma ray pulsars in the *Fermi Gamma Ray Space Telescope* dataset was performed on the *Pleiades* cluster and resulted in the cover page of *Science* (1).

Krumholz's research group studies star formation and the interstellar medium (ISM), and uses simulations on *Pleiades* to answer basic questions such as what sets the rate of star formation in galaxies, what determines the stellar initial mass function (IMF),

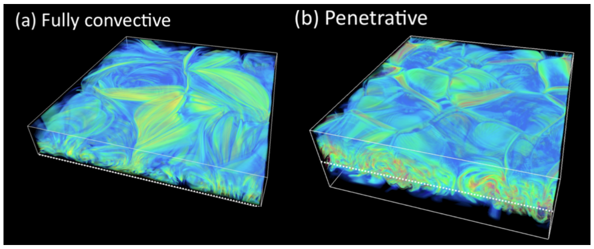

Figure 5: Dynamos in non-penetrating and penetrating convection. Shown are volume renderings of the magnetic field (with strong field shown in bright and opaque colours, weak as darker, translucent colours) generated by simulations of both confined convection and convection overlying a stable region (penetrative convection). (From (11).)

and what sort of interstellar environment produced the Sun. Krumholz and undergraduate L. Gendelev performed radiation magnetohydrodynamic simulations using the ATHENA code to study how star-forming clouds are disrupted by the ionizing radiation of newly-formed stars (50, Fig. 6). Undergraduates D. Dukes and J. Craig have used *Pleiades* to simulate N-body interactions between stars in the Sun's birth cluster, which might have helped shape the architecture of the young Solar System (33). Graduate student N. Goldbaum is using *Pleiades* to simulate both the evolution of individual star-forming molecular clouds, and populations of clouds in a galaxy, using the ENZO adaptive mesh refinement (AMR) hydrodynamics code. Graduate students M. Fumagalli and R. da Silva have worked with Krumholz using *Pleiades* to develop the stochastic stellar population synthesis code Stochastically Lighting Up Galaxies (SLUG) for the study of the statistical properties of star formation in dwarf galaxies with low star formation rates (40; 24). Krumholz worked with ISIMA student M.-K. Lin using the ORION AMR radiation-hydrodynamics code to study what sets the rotation rates of massive stars at birth (118), and with ISIMA student A. Myers to study how metallicity affects the fragmentation of interstellar clouds and the resulting IMF (134). Postdoc C.-C. Yang has used simulations with the PENCIL code on *Pleiades* to study how metals are mixed by turbulence in the ISM, and then incorporated this into stars, producing the chemical abundance patterns that we observe.

Planetary Systems Lin and collaborators (59) used *Pleiades* to simulate the supernova triggered collapse of a proto-solar molecular cloud. Their results demonstrated for the first time that adequate amount of radioactive ^{26}Al isotope may be intercepted and the solar nebula can form on a time scale shorter than that recently inferred from the calcium-aluminum rich inclusions age spread in chondritic meteorites. The discovery of diverse dynamical architecture in extra solar planetary systems clearly indicate that during their formation, some protoplanets tidally interact and migrate extensively through their natal disks. The first systematic studies of this process in 2D turbulent disks was carried out by (8). Their results naturally account for the retention of a large population of newly discovered super-Earths around nearby stars.

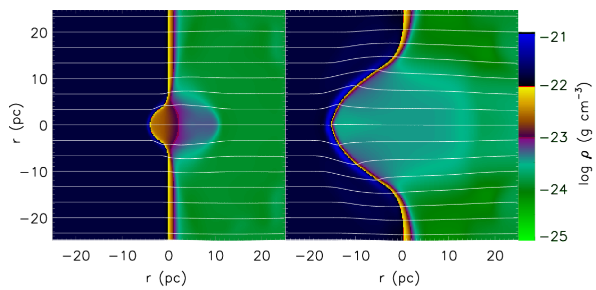

Figure 6: Two time slices from a simulation of the evolution of a blister-type H II region in a magnetized medium. Color indicates gas density; lines are magnetic field lines. In this simulation a star placed near the edge of a dense molecular cloud in the left half of the simulation box begins to emit ionizing photons, which heat the gas and drive a shock into the cloud on the left, coupled with a jet of ionized material rocketing away to the right. This simulation was carried out on *Pleiades* as part of Leo Gendelev's senior thesis research (50).

Nimmo's work using *Pleiades* has focused on three areas: 1) the origin of the Martian hemispheric dichotomy, where its formation was modeled via a giant impact using the Zeus hydro code (138). It was shown that a narrow range of impact sizes/velocities could generate a basin of the required magnitude, without producing so much melt that the basin was erased; 2) an N-body integrator to investigate the orbital evolution of satellites in the presence of tidal torques. This integrator has been used to place constraints on the interior structures and thermal histories of Enceladus, Dione and Tethys (204); and 3) a 3D spherical convection code to investigate tidal dissipation within and thermal evolution and relaxation of various icy satellites, including Enceladus

(158; 159), Iapetus (163), and Pluto (162). UCSC graduate student E. Chen, advised by Nimmo and Glatzmaier, has run 3D simulations on *Pleiades* of the circulation in the subsurface ocean on Europa and the resulting magnetic induction due to Jupiter's time-dependent magnetic field through which Europa orbits. Glatzmaier has run 3D simulations on *Pleiades* of the convective dynamos of giant planets, like Jupiter and Saturn (see Fig. 7), and of terrestrial planets, like the Earth and Mercury.

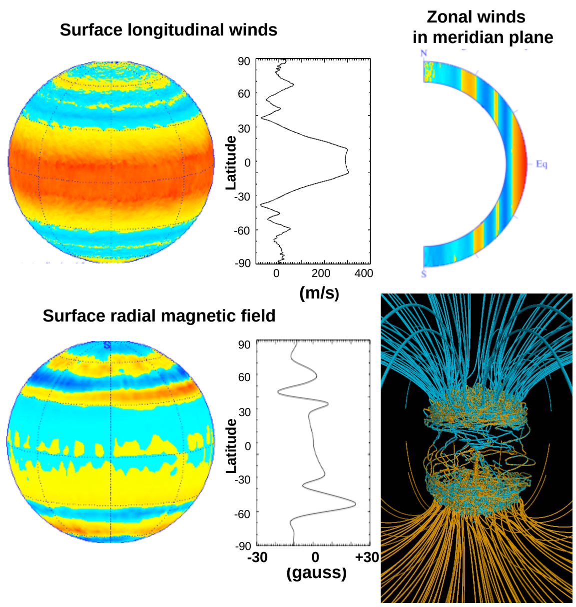

Figure 7: A snapshot of a Saturn dynamo simulation showing the structure of the flows in the upper row of the figure and the generated magnetic field in the lower row.

Laughlin and collaborators have used *Pleiades* to demonstrate, for the first time, that the eight planets of the solar system can exhibit a long-term instability in which Jupiter and Mercury evolve into secular resonance (10). Using a long-duration N-body simulation of the solar system evolution, they found that when the secular resonance is achieved, Mercury's orbital eccentricity increases until its aphelion position intersects Venus' orbit. At this point, the inner solar system becomes dynamically unstable and various catastrophes can ensue. The simulation showed that there is a finite (albeit small) chance of this outcome during the next 5 billion years. They also made extensive use of the cluster to compute global climate models for extrasolar planets on highly eccentric orbits. Comparison of such models with *Spitzer* observations of the transiting planet HD 80606b showed that the radiative timescale in the planetary atmosphere is surprisingly short, of order 12 hours (as compared with 3 days on Earth) indicating the presence of an effective absorbing constituent high in the planet's atmosphere. These results were published in *Nature* (111) and the ray-tracing simulation of the optical appearance of the planet (also done with *Pleiades*) was featured on the cover of *Nature*.

Fortney has been investigating the 3D structure and circulation of exoplanet atmospheres. The "hot Jupiters" are a class of strongly irradiated, tidally locked planets, that orbit in only a few days. Fortney has co-developed a model, SPARC, which simultaneously solves the 3D dynamics along with 1D non-gray radiative transfer. The code has been run on *Pleiades* and used to investigate the dynamics, emission spectra, and transmission spectra of several well-characterized planets (173; 174; 115; 39).

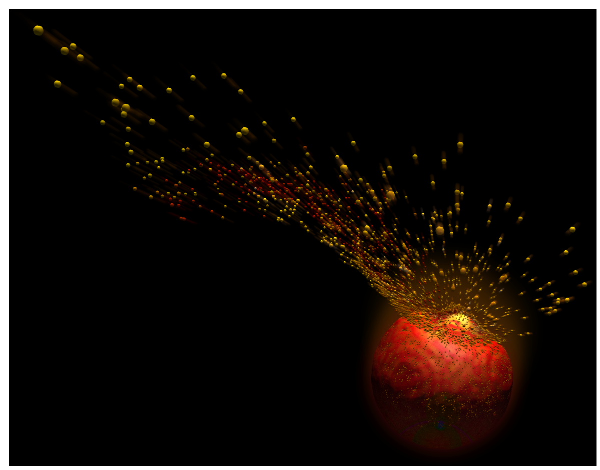

Figure 8: Simulation of the formation of the northern lowlands of Mars. This collaboration (129), including 3D SPH simulations performed on *Pleiades*, showed that a reasonably oblique impact (45°) would produce a crater with low melt production and form an ellipsoidal mega-crater matching the shape of the dichotomy boundary. This figure was the cover of *Nature*.

Asphaug's 3D hydro code models of the Mars hemispheric dichotomy (129; 130) established that a major collision at 30-60° off axis, an event which can only be modeled in 3D, allows for a mega-crater to form that is of the right shape and that is formed with only limited production of melt due to the off-axis nature of the collision (see Fig. 8). Graduate student L. Chambers (15) modeled the UV emissions resulting from hypervelocity impacts (30-50 km/s) into Saturn's ring particles, while graduate student L. Ong (148) performed supercomputer models of comets impacting the Moon, to derive the fraction of water ice that remains bound and make quantitative estimates of the contribution of cometary water over the past 2 billion years. Simulations, run on *Pleiades*, of planetesimal collisions during the first few million years of solar system history, have led to a new model for the origin of chondrules, the igneous melt droplets found in profusion in chondrite meteorites (7). *Pleiades* has also enabled geologically realistic models of planetary collisions: (78) ran supercomputer models to predict the geologic setting that NASA's Dawn spacecraft would observe when it entered rendezvous about asteroid Vesta, predicting the diameter and topographic profile of the mega-crater with unexpected accuracy. (77) modeled the origin of the lunar farside highlands as an accretionary pile of material added to the Moon, when it accreted a companion moon about 1/3 its diameter (Fig. 9). This modeling prediction, featured in magazines and newspapers worldwide, will be tested by NASA's GRAIL mission.

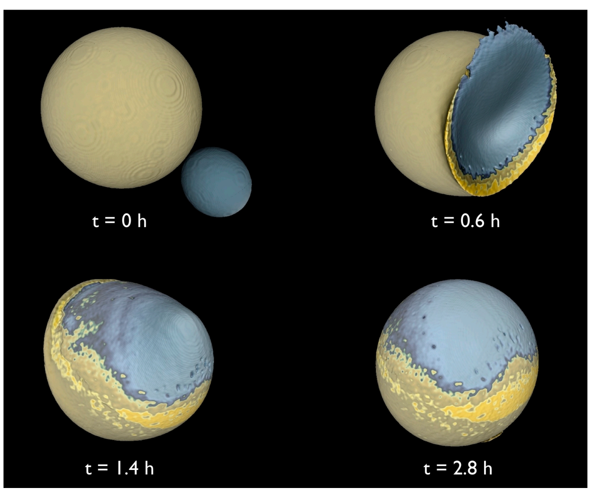

Figure 9: Simulation of the Moon's farside highlands by a giant 'splat' involving a second moon, according to the hypothesis of (77). This simulation was run on *Pleiades* and visualized on its graphics front end.

2.2.3 Proposed New Research in Astrophysics

The presence of *Pleiades* for the last five years has helped to grow a world-class program in computational astrophysics. For that program to endure and expand, it is now essential to build a new HPC facility. Here we discuss some of the science that would be enabled by that machine. The effort will involve a total of at least 16 faculty (we can reasonably expect a few new hires over the coming years), and over 80 postdocs and graduate students. This lower limit is less than past usage, and will be easily attained if each faculty currently associated with this proposal supervised only 5 graduate students and postdocs in the next five years. Many are supervising more than that right now.

Galaxies and Cosmology Madau's research over the next several years will focus on numerical studies that will lead to an understanding of the building block of galaxies, of their star formation activity at high redshifts, of the enrichment of the circumgalactic medium (172), of the formation of realistic late-type spiral galaxies (60), and of the processes that determine the survival of substructure in present-day galaxy halos, all in the standard cold dark matter paradigm of structure formation in the universe. He will continue his program of cosmological hydro AMR simulations with the ENZO code (http://code.google.com/p/enzo/) to address in a coherent way the formation of dwarf galaxies with H_2-regulated star formation at high redshift (104), their contribution to the ionizing background that ionized the universe (214), and their detectability by the *James Webb Space Telescope*. Many of today's "observables" within the Milky Way and nearby galaxies relate to events occurring at early epochs, during and soon after the era of reionization. Madau will use a combination of ultra-high resolution N-body (with the code PKDGRAV2, which has demonstrated good scaling on *Jaguar* at ORNL on over 12,000 CPU) and SPH simulations (with the code GASOLINE (222)) to provide a better characterization of the assembly of the Milky Way halo from such subunits, of the small-scale structure of Galactic dark matter, of the "missing satellite problem", and of the abundance and properties of stellar streams and debris that are the signature of disrupted satellite galaxies.

Primack's continuing research program in this area will focus on running a large set of cosmological hydrodynamic simulations of galaxies using the ART adaptive mesh code, in support of

CANDELS and other observations. CANDELS (http://candels.ucolick.org) is the largest project in the history of the *Hubble Space Telescope*, and will take three years to complete. The simulation program will include radiative transfer calculations with the SUNRISE code of the appearance and spectral energy distributions of galaxies including stellar evolution and dust scattering, absorption, and reemission of radiation. These simulations are currently at 15-30 parsec resolution, and include previously neglected phenomena such as radiative feedback. The latest NVIDIA GPUs on the new proposed machine will speed up SUNRISE calculations by an order of magnitude. Primack will continue to run his *Bolshoi* program of large dissipationless cosmological simulations on remote supercomputers. *MiniBolshoi* will include the evolution of hundreds of Milky Way size galaxies and their subhalo populations, providing poorer spatial and mass resolution than simulations of individual halos such as *Aquarius* and *Via Lactea*, but with far greater statistics. The proposed new HPC cluster will be used to analyze the large simulation outputs and compare them with observations.

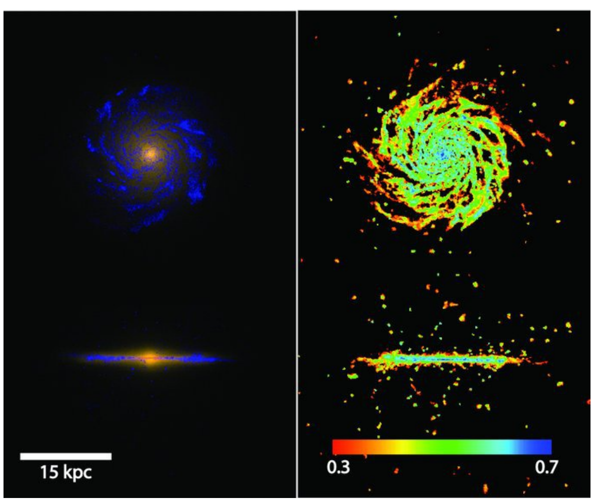

Figure 10: "Eris", the first cosmological N-body/SPH simulation in which a close analog of a Milky Way disk galaxy arises naturally (60). Left panel: the optical/UV stellar properties of Eris at $z = 0$. The images, created by running the radiative transfer code SUNRISE on *Pleiades*, show an i, V, and FUV stellar composite of the simulated galaxy seen face-on and edge-on. Right panel: projected face-on and edge-on surface density maps of Eris's neutral gas. The color bar shows the neutral gas fraction.

New UCSC faculty C. Conroy is developing a new generation of stellar interior, atmosphere, and spectral models for arbitrary elemental abundance patterns. These models will be used to measure the abundance patterns, ages, and stellar initial mass functions of distant galaxies and of Galactic globular clusters. This work is utilizing the new modular stellar evolution code MESA and the ATLAS12/SYNTHE routines developed by R. Kurucz. Uncertainties in the atomic and molecular linelists will be reduced by running an ancillary grid with variations in the linelist parameters that will be compared to high resolution spectral standards such as the Sun and Arcturus. On another front, Conroy's group is developing new modules for the ENZO AMR hydrodynamics code in order to study the interaction of AGB winds, Lyman-Werner flux from B stars, and supernovae on the development and retention of gaseous reservoirs in young globular clusters. Galactic clusters appear to have undergone multiple generations of star formation at early times. This approach will be the first to include all the necessary physics and processes in order to probe the origin of multiple stellar populations in globular clusters.

High Energy Astrophysics One of the major research thrusts by Woosley's group in next several years will be models for Type Ia supernovae (SN Ia) using the CASTRO (212) and MAESTRO (219) codes. CASTRO is a 3D compressible hydro code with AMR that has been developed chiefly with the SN Ia application in mind. MAESTRO is a low-Mach number code, also 3D AMR, optimized to study the pre-explosive convection that sets up the ignition conditions for thermonuclear runaways in stars. MAESTRO can take time steps that are not Courant limited and works very stably in environments where small driving forces operate in the presence of an otherwise tight hydrostatic equilibrium. Both codes have demonstrated good scaling on *Franklin* at NERSC and *Jaguar* at ORNL on over 100,000 CPU. Over 60 M CPU h has been awarded in 2012 alone to run SN Ia models with these codes at ORNL, NERSC and on the NSF *Blue Waters* Cray. Substantial local developmental work will be necessary, however, in order to effectively utilize these resources. This local work falls in three areas: 1) developing and testing efficient algorithms and approximations for treating nuclear burning, especially detonation physics in CASTRO and flame physics in MAESTRO; 2) 2D surveys of several classes of models for SN Ia, especially Chandrasekhar mass models and sub-Chandrasekhar mass models; and 3) improving and testing the performance of both codes on GPUs. The latter is especially important since the new machines at ORNL and *Blue Waters* will make extensive use of GPUs. With MAESTRO it will be possible to study, for the first time, the convective runaway that goes on for months before the thick helium shell in the sub-Chandra white dwarf finally runs away, allegedly culminating in a detonation. Such studies will show the localization of the burning and whether ignition occurs in one point or many and whether it occurs at the base of the accreted layer or somewhere above. In other studies, Woosley will explore multi-dimensional models for Type I X-ray bursts on neutron stars (215) and the formation of proto-neutron stars in core-collapse supernovae. X-ray bursts will be calculated using MAESTRO, but tables will have to be developed to approximate the complex burning of hydrogen by the ”rp-process”.

Ramirez-Ruiz’s group will be developing an SPH general relativistic framework to study the mergers of compact stars and black holes. More realistic models for tidal disruption of stars by massive black holes will be implemented in the FLASH code. Detailed simulations will tell us what happen when stars of different types get tidally disrupted, and what is the radiation spectrum and light curve a distant observer may detect as the observational signature of such events. Ramirez-Ruiz will also be developing a 3D a general relativistic framework to study the evolution of neutrino-cooled disks using the HARM3D code.

Stars and the Sun Brummell and Garaud will continue the groundbreaking theoretical work based on basic fluid dynamical simulations using *Pleiades* that are currently under way. The search for a dynamo scenario that operates more efficiently than the standard kinematic model at high magnetic Reynolds numbers will continue. Essentially nonlinear dynamos, where the dynamo-generating flows are driven by magnetic forces themselves, are showing great promise, but such simulations with HPS are necessarily high resolution and the process is highly

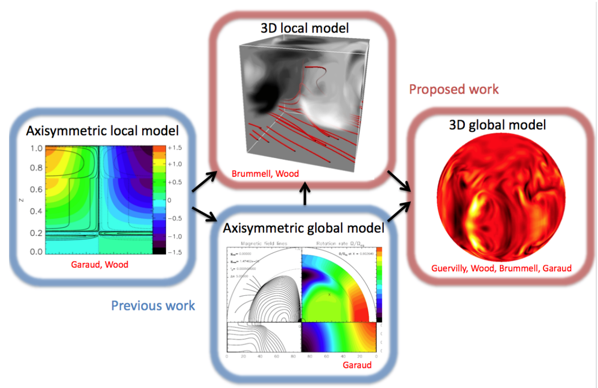

Figure 11: Interaction of many different models towards an understanding of the solar tachocline by Brummell and Garaud’s GAFD group.

nonlinear and therefore temperamental, requiring extensive preliminary searches of a wide parameter space before final production simulations. The new machine proposed in this document would be ideal for this task.

The group’s work on the dynamics of the solar tachocline has finally revealed why previous numerical simulations have not exhibited dynamics that are quantitatively similar to the observed phenomenon. Preliminary simulations of a new approach on *Pleiades* using HPS have shown great promise in a simplified Cartesian geometry and this theory now needs extensive exploration in the proper geometry using the PARODY code.

Garaud and collaborators will make their extensive dataset of existing numerical simulations on double-diffusive convection available to the general community through the proposed new machine. In addition, they plan to begin running larger-scale simulations to characterize transport by layered convection in the parameter regime relevant for giant planets and stellar interiors and to study the generation of large-scale internal waves and their interaction with the turbulence. This is of interest in order to explore a new mechanism for the excitation of stellar pulsations (for comparison with Kepler observations) and to determine whether layer formation could be triggered mechanically in systems which are not otherwise unstable to layering.

Glatzmaier and collaborators will be benchmarking their new solar dynamo code against Glatzmaier's older code, to compare the different formulations and the use of finite-differences versus Chebyshev polynominal expansions in radius. After benchmarking, high resolution, highly parallelized simulations will be run of a complete solar interior, resolving the shear and magnetically-driven instabilities of the tachocline, magnetic field generation in the convection zone and magneto-inertial-gravity waves in the radiative interior.

Krumholz's group is now developing new simulation techniques to understand star formation over galactic and cosmological scales. Postdoc J.-H. Kim is leading work to use ENZO to simulate the star formation within an entire galaxy including ionizing radiation transfer, using AMR to simultaneously resolve individual star-forming clouds. This will allow for the first time a simultaneous determination of the star formation rate, the properties of molecular clouds, and the ionizing photon escape fraction from galaxies. Graduate student N. Goldbaum is developing new, much more realistic subgrid models for star formation in cosmological simulations, that will be implemented in the ENZO code. Graduate student A. Rosen will run prototype simulations of the formation of super star clusters and globular clusters including dust-reprocessed radiation feedback using the ORION code. These simulations should for the first time yield realistic answers to basic questions about globular cluster formation, such as whether the globular clusters we see today are only a small remnant of those that formed, as is the case for smaller open star clusters, or whether globular clusters form with such high gas-to-star conversion efficiencies that most survive.

Planetary Systems Short-period gas giant planets are exposed to intense radiation from their host stars. Dobbs-Dixon, Lin, & Wang have constructed a fully 3D radiative hydrodynamic code to simulate both the dynamics and CO2/CH4 cycle in the atmosphere of these hot Jupiters. Their preliminary results showed that non equilibrium reaction rates can lead to strong day/night intensity contrasts which may be observable with the *Spitzer* space telescope. Their next task is to considered the water cycle on super-Earth planets. Liu, Guillochon, Lin, Ramirez-Ruiz have implemented an equation of state for gas giant planets' cores which will enable them to extend their simulations of planetary tidal disruption (64). They will investigate the possibility of retaining gas giants' cores after the tidal disruption of their envelope. Li and Lin will implant radiative diffusion into their hydrodynamic code and simulate gas accretion onto supermass cores in protostellar disks.

Future work by F. Nimmo will focus on two areas. First, satellite evolution is a coupled problem: thermal evolution affects orbital evolution, and vice versa and so a N-body orbital code will be combined with a 3D convection code to investigate satellite evolution in a fully self-consistent manner (164). This code may be applied not only to outer solar system bodies, but also to the Earth's Moon, a focus of much current scientific interest. Second, Nimmo will continue to use his 3D spherical conduction code to model satellite thermal evolution. Particular areas of interest will be to determine whether Enceladus may have undergone episodic convection which could explain its puzzling heat output, and the extent to which his models can be reconciled with newly-derived observational constraints on satellite thermal history from basin relaxation and flexural studies.

Glatzmaier plans to continue his simulation studies of Europa's ocean circulation and magnetic induction, of the laboratory dynamo experiment and of giant-planet and terrestrial planet dynamos using more processors on the new cluster, which will allow greater spatial resolution. In particular, the simulations of Mercury's dynamo will help to explain the recent magnetic field measurements made by NASA's Messenger Mission to Mercury and the dynamo simulations of Jupiter and Saturn will make predictions for the discoveries that will be made starting in 2016 by NASA's Cassini Solstice Mission at Saturn and its Juno Mission to Jupiter.

Laughlin's group has begun obtaining radial velocity data from the *Automated Planet Finder* telescope, which can probe the terrestrial-mass planetary census of nearby stars. Large suites of dynamical stability calculations are required in order verify system stability and resonant structure within the emerging systems. In addition, he has begun to work with MHD models of planetary atmospheres for planets that oscillate in and out of the regime where the planetary magnetic field couples with the atmospheric dynamics. Studying these systems, even with certain idealizations, is numerically intensive, requires many exploratory simulations, and is ideally suited to an in-house resource such as *Pleiades*' envisioned successor.

Fortney will continue to model the dynamics and spectra of the highly irradiated transiting planets, but also move into other areas. The atmospheres of low-mass planets (below 10 Earth masses) are now being characterized. These atmospheres may be made up of predominantly of hydrogen, nitrogen, or carbon dioxide. His group will model the atmospheric dynamics and spectra of these planets, which are being found by Kepler and other platforms. In some cases, for Earth-mass planets, he will assess the habitability the planetary atmosphere/surface interface.

Asphaug's group will continue working on pairwise accretion, and the computed physics and petrology of planets, planetary remnants, and meteorites. Graduate student N. Movshovitz has begun the task of GPU acceleration of the parallel SPH code in preparation for next generation modeling. Graduate student J. Guillochon has a NASA Fellowship to model the Moon-forming giant impact. The group will develop hit and run models for the origin of planet Mercury, now a great puzzle since the reigning explanation for its oversized iron core, as the aftermath of an extraordinarily energetic giant impact, is contrary to measurements by the MESSENGER orbiter indicating little volatile fractionation. The group shall continue studying smaller collisions involving asteroids, comets, and the earliest meteorite-forming planetesimals, as well as Moon-forming mergers around Jupiter and Saturn, and binary collisions in accreting exoplanetary systems to help characterize these newly discovered populations.

2.2.4 I/O and Storage Management

While computational scientists make heavy use of the I/O and storage system, it is often difficult to instrument code running large-scale clusters to analyze and improve I/O performance. The computer scientists in our group (Miller and Brandt) will work with the computational scientists to explore both short-term (seconds to hours) and long-term (days to months) I/O access patterns of code running on the new cluster. This analysis will improve our understanding of how cutting-edge astrophysics codes use large-scale clusters, and will also provide insight on how scientists find and utilize data. At UCSC, we have extensive experience with scientific file systems and data access patterns (223; 211); we will instrument system to allow us to gather additional I/O traces from real-world applications without burdening applications or scientists. Because this system will be in an open environment, we will be able to make these traces available externally, providing a valuable resource for high performance computer science researchers. We will also gather information on metadata usage—*which* files and data the applications and scientists use. We are already exploring the problem of improving file naming systems for high-performance computing (217; 218; 220); the unique opportunity for collaboration between computer scientists and computational scientists in an open environment will help guide our development of new naming systems to make large-scale storage easier to use.

In addition to short-term usage studies, we will explore long-term usage patterns for scientific data (208; 210) by transparently monitoring accesses to the petabyte-scale archive of scientific data. This will not disturb users of the system, but will lead to better understanding of long-term scientific data usage on a publicly-accessible petabyte-scale archive, helping us design more efficient, more reliable archival storage systems (209; 224) and leading to reduced long-term data storage costs across a wide range of scientific disciplines. As with short-term tracing, we plan to make our long-term access traces available publicly, providing a much-needed resource to the computer systems research community.

2.3 Impact on Research and Training Infrastructure

In addition to amplifying the productivity and outside funding of senior investigators, the new cluster will have a major impact on the training of graduate students and postdocs. Computational science and simulation are increasingly playing key roles in research. A large fraction of the science done in the next decade will be carried out on mid-size and large computer clusters by the people that have the special training to use them effectively. This training is best achieved in graduate school and as a postdoc by a "hands-on" approach. The best evidence of our success in this area are the large number of publications where the lead author is a graduate student or a postdoc who used our previous machine (see References Cited). Tens of such students have gone on to prestigious postdoctoral positions. More than a dozen of our computationally-oriented students and postdocs (Fabrycky, Chen, Diemand, Kasen, Langton, Ogden, Porter, Robinson, Roepke, Rogers, Stellmach, Yu, Zemp, Zingale) have taken faculty jobs in theoretical astrophysics. The field of computational astrophysics is presently heavily male-dominated. And while only one senior member of our team is a woman, we are actively trying to improve gender balance: over the past 5 years, 16 female students (Lindsay Chambers, Erinna Chen, Qiaoning Chen, Jennifer Holt, Candace Joggerst, Elizabeth Lovegrove, Jill Naiman, Darcy Ogden, Lissa Ong, Tami Rogers, Sarah Robinson, Anna Rosen, Erica Rosenblum, Adrienne Traxler, Katelyn White, Angie Wolfgang) and 4 postdocs (Celine Guervilly, Annalisa Pillepich, Sijing Shen, Quingjuan Yu) have been given access and used *Pleiades* to pursue their research. The following UCSC undergraduate students, also *Pleiades* users, have been awarded graduate student fellowships: Luke Kelley (Harvard), Matt Coleman (UCSB), Ricky Fernandez (Columbia), and Rion Parsons (Cambridge). We anticipate that the success of this hands-on training program will continue with new resources.

Our experience has also shown that the proposed computing platform will facilitate faculty recruitment in theoretical astrophysics. Over the last 5 years, UCSC's excellent computational resources have helped attract 5 new theoretical astrophysicists and planetary scientists to the faculty (Nic Brummell, Charlie Conroy, Jonathan Fortney, Mark Krumholz, and Francis Nimmo). Having a state-of-the-art HPC cluster to share and work on attract new hires in the first place and then cata;yses collaborations with the more senior faculty, increasing the collective productivity and fostering interdisciplinary collaborations.

UCSC is host to a number of institutes for which the proposed system would play a role. UCSC runs the University of California systemwide "High Performance AstroComputing Center" (HiPACC, http://hipacc.ucsc.edu), which links computational astrophysicists across the UC campuses and the associated National Laboratories. Each year HiPACC sponsors conferences, public outreach via a website with news stories, media releases including video visualizations, planetarium shows, grants to promote collaboration between campuses and the associated National Laboratories, grants to support undergraduate research on computational astrophysics, and an international summer school for graduate students and postdocs. The summer schools in 2010 on "Hydrodynamic Galaxy Simulations" (at UCSC) and in 2011 on Computational Explosive Astrophysics (at UCB/LBNL) had as students many of the best young computational astrophysicists in the world. The 2012 school will be on "AstroInformatics" (at UCSD/SDSC), and the 2013 school will be on "Star and Planet Formation" (at UCSC). The proposed HPC cluster will provide support to HiPACC scientific, educational, and outreach efforts. These include enabling code development and research, preparing and hosting simulation outputs and related astronomical data, and preparing images and visualizations from astrophysical observations and simulations for outreach and education. UCSC also hosts the "Next Generation Telescope Science Institute" (NEXSI, http://nexsi.ucolick.org), which promotes computational research in support of new multi-billion dollar astronomical facilities, and "Theoretical Astrophysics at Santa Cruz" (TASC), which has begun a new interdisciplinary program in scientific computation and visualization in collaboration with members of UCSC's Digital Arts and New Media Center. UCSC has also launched in 2010 the International Summer Institute for Modeling in Astrophysics (ISIMA), a unique visitor program which combines graduate training with collaborative research on outstanding problems in astrophysics, using mainly local computing resources. Outside of astrophysics, the UCSC Computer Sciences Department is also a leading center for research into the use of high-performance storage in support of large-scale scientific computing. Two groups within the School of Engineering have a history of highly visibile successes in the field, including the development of the *Ceph* file system, investigation of real-time storage for large-scale computing, and analysis and development of systems for long-term storage of scientific data, highly pertinent issues not merely confined to astrophysical research. It is the hope of the Co-I's of this proposal that the new HPC platform proposed here will further cement UCSC's position as a center for computational astrophysics and a portal for simulation results and educational resources.

Indeed, one of our goals is the nurturing of a new kind of interdisciplinary computational scientist, well versed in both astrophysics and the computational tools of large-scale simulation and analysis. Computational power has increased enormously in recent years, but educating people to use it has lagged. Few graduate programs currently aim specifically at large-scale computing. Our team will establish core courses in High-Performance Computing at UCSC, assisted by course development in the Applied Math department. The courses will combine computer-science topics (e.g., applied math, algorithm design, parallel programming, data analysis, visualization) with relevant physics and astrophysics, plus hands-on problems solving using the proposed resources. The guiding principles for the graduate program are based on recommendations in "Re-Envisioning the PhD" (http://www.grad.washington.edu/envision). UCSC's rich environment naturally provides novel opportunities through multiple mentors, exposure to a wide variety of career options, interdisciplinary activities, training for diverse teaching options, partnerships with institutions that prepare and hire doctoral students, and alliances with existing programs at partner institutions for recruiting and retaining women and underrepresented students. As high-performance computing becomes a routine tool, industry and government will seek graduates with expertise in this field. The broad palette of astrophysical simulations makes them an ideal training ground for many job situations. Students and postdocs throughout our team will have access to internships via an extensive network of partnerships with industry and national laboratories, including Google, NASA Ames, LBNL, and LLNL.

Our team will also conduct an aggressive knowledge transfer program, providing the broader public the opportunity to do and see astrophysics. The visual appeal of astrophysics simulations will be used as a tool to engage students in research, connect communities as disparate as astrophysics and the visual arts, and disseminate the scientific messages of the Center via partnerships with two distinguished high schools (Harker Academy and Pacific Collegiate Charter School), planetaria, the NASA Ames Hyperwall, GalaxyZoo, and Google Sky. In a close partnership with UCSC's MFA in Digital Arts and New Media and UCSC's top-ranked program in Science Communications, arts students and participating faculty are currently collaborating with computational students on joint visualization projects through our Openlab initiative (http://artsresearch.ucsc.edu/openlab/). This basic design facilitates diversity by linking two very different communities, applied computational science and digital arts. Including the latter field, better balanced with regard to gender and ethnicity, will immediately create a more diverse and vibrant student body. Engaging students in exciting projects that link these two communities will further strengthen these bonds.

2.4 Description of the Research Instrumentation and Needs

2.4.1 Instrumentation

It is beneficial to first review the configuration of our current cluster *Pleiades*, which will be used as the reference point for the proposed new computing environment. *Pleiades* includes 208 compute nodes, each with 2 dual-core Intel Xeon processors and 8 GB memory (2GB per core), and 2 visualization nodes, one with 2 dual-core Intel Xeon processors and 32 GB memory, and the other with 2 quad-core Intel Xeon processors, 64 GB memory and a Nvidia Quadro FX 5600 graphics card. It has 55 TB of total usable storage (70 TB raw), served from 4 file servers using the IBRIX parallel file system. The cluster is interconnected with 2 network fabrics: Gigabit Ethernet and non-blocking DDR (20 Gb/s) InfiniBand.

Many of the design principles for *Pleiades* still hold true for the proposed new cluster. The diverse needs of our research groups still demand a general-use, adequate memory (2 GB/core), and fast internetworks machine with sufficient storage served from an efficient parallel file system. On the other hand, the specifications of the proposed new cluster also reflect the lessons we have learned from using *Pleiades*, the shift of parallel programming models (including, among others, the advent of the multicore and manycore computing), and the new data-intensive paradigm that is transforming scientific computing and the way we do Astronomy. In addition, we will also strive to retain and reuse any components of the *Pleiades* cluster (like racks, UPS, power and cooling system, visualization nodes, etc) whenever applicable.

The choice of hardware is set by a balance among our needs, cost estimates, and siting issues. An on-campus facility with a few thousand processing units is ideal for code development, running small- and medium-scale experimental simulations, data analysis, and student/postdoc training. Working with vendors, we have determined that while the computing power of a state-of-the-art compute node doubles every 18 months, the price and power (and cooling) requirement of a node stay roughly constant (which is a logical conclusion of Moore's law). Without spending capital for upgrading the power and cooling capacity of the data center at UCSC, we are limited to around 200-250 compute nodes.

Quotes have been solicited from four vendors - ASA, Dell, Fine Tec, and SGI. The quote from Dell is highlighted here for multiple reasons: 1) Dell's quote provided the greatest detail; 2) Dell was the vendor for the *Pleiades* cluster and has provided excellent products and services; 3) Dell's price is competitive; 4) using Dell to purchase the new cluster would make it easier to retain and reuse some of *Pleiades'* components. All quotes are given in the Supplementary Documents section and are (with the exception of SGI) very similar in total cost and specifications, giving us confidence that the estimated hardware budget, $1.04 M, is realistic given our requirements. The actual vendor and quote will only be selected after extensive product review and interviews once the proposal has been funded. As the new machine would most likely be assembled from components available in late 2012, it will probably differ in details from what follows. In particular, we anticipate the availability of faster processors with more cores, faster GPU units, and higher capacity mass storage solutions for the same cost. The performance increase over *Pleiades* is expected to be a factor of about 3 to 10, though this factor is code and application dependent. Two factors will contribute towards the expected performance boost: 1) there will be 3 times the number of CPUs cores (6 times if hyperthreading is enabled), with a small boost in efficiency per core over *Pleiades*; 2) a fraction of compute nodes (10 to 20 percent) will be GPU accelerated, which can provide a 10 to 100 times speed increase for codes that are optimized for GPU computing.

2.4.2 Cluster Interconnects

Infiniband is the de facto standard for interconnects of HPC clusters, with its extremely low latency and high bandwidth. We'll use a non-blocking, QDR InfiniBand fabric for the proposed cluster, which will provides a theoretical bandwidth of 40 Gb/s between any two nodes in the cluster, with a latency of a few microseconds. This is an upgrade from the DDR (20 Gb/s) InfiniBand fabric of the *Pleiades* cluster. Additionally, the proposed cluster will also have a Gigabit Ethernet network, mostly for management traffic. We will recycle the components from the *Pleiades* cluster, thus saving some costs.

2.4.3 Compute Nodes

Given the contraints on power and cooling capacity of our site, the proposed cluster will be limited to 200-250 compute nodes. Regarding CPUs, our experience with many small to mid-size clusters, as well as the benchmarks independently done by many hardware review sites, has led us to focus on the Intel Xeon product line, which holds the performance crown in terms of both total performance and performance per watt, particularly against the AMD Opeteron line. Each compute node will consist of two hex-core Intel Xeon X5650 CPUs, thus providing 12 cores per nodes (24 cores if hyperthreading is enabled).

There will be 24 GB memory per compute node (2 GB per core) – all of our applications will run comfortably with that memory configuaration. Each node will include a small hard drive. Their low cost (less than $20K for all nodes) provides the flexibility of a local swap space or local scratch storage for future applications. Between 10 and 20 percent of the compute nodes will be GPU accelerated. The Dell quote includes 20 such accelerated nodes, each with an nVidia Tesla M2070 GPU computing unit. GPU computing nodes will serve multiple goals. They will make a powerful subcluster for visualization and data analysis. They will also help our students and postdocs become familiar with the increasingly important art of GPU and manycore computing, and prepare them for the petaflop-scale GPU supercomputers at national centers, including *Blue Waters* at NCSA and *Titan* at ORNL.

2.5 Management Plan

2.5.1 The Site

The timing of this proposal is largely motivated by the declining ability of our current cluster, *Pleiades*, which has been serving the computational astrophysics community at UCSC for more than 5 years. We will not, however, build a cluster from scratch, as in the case for *Pleiades*. The new cluster will rather be a major upgrade of *Pleiades*. We plan to use the same site, and retain and reuse any component whenever applicable of the old cluster.

Pleiades is hosted at the Data Center of Information Technology Services (ITS) at UCSC. In order to deploy *Pleiades* 5 years ago, UCSC spent about $450 K to renovate the site, adding a capacity of 100 kW power and 25 tons cooling, and a floor space for 8 racks. The infrastructure is in perfect working order, and thus makes the best home for the proposed machine. As the letter from the Vice Chancellor for Research (see Supplementary Documents) attests, the campus is committed to providing the site and power/cooling for the proposed machine, if this proposal is funded.

The proposed machine will fit nicely within the power, cooling, and space budgets of the site. One additional benefit of hosting the new cluster at the ITS data center is the availability of high-speed dark fiber connections. UCSC has recently deployed Dark Fiber infrastructure, providing 10 Gbps layer 3 circuit to CENIC and thus a superfast pipe to other major reseach universities and national labs. We will link the proposed machine and the storage subsystem directly to the Dark Fiber network. This will allow large simulation data to be moved from national facilities, and to become accessible by the community at large.

2.5.2 Machine Management and Time Allocation

Through the previous MRI-funded machines, UCSC has accumulated invaluable experience and expertise in cluster management and operation. We plan to use the open source Rocks Cluster Distribution, which is developed by SDSC. We'll likely purchase a support contract for the Lustre file system, a license for Intel compilers and Intel cluster suite, and maybe a license for a parallel debugger, either Totalview or DDT.

Hardware is warrantied for 3 years in the quotes from the vendors. During final purchasing negotiations we aim to secure an advantageous bid on warranty support in the extended years (4th and 5th years).

The machine will be operated by Dr. Shawfeng Dong, the current administrator of *Pleiades*. Dr. Dong is one of the top experts in parallel computing at UCSC, and excels in both the hardware and software aspects of high performance computing. In most national supercomputing centers, there are two types of supporting staff: system adminstrators (who keep the computers up and running) and consults (who help users run their parallel codes); at UCSC, Dong has done a superb job filling both roles. While this proposal is for 2 years of operation, we expect the useful life of the machine to be five or more years. During the out years, the non-hardware costs of operations will be paid for by other grants or donors.

Time allocation on the computer will be overseen by a committee consisting of the PI, the 4 co-PIs, and the 11 Senior Associates on this proposal. Most of the members of our team are users of *Pleiades*, and have shared such resource equitably in the past. Computer time will be allocated equally among all the members of our team. Time will be measured on average usage over many months; there will certainly be periods when one or a few users dominate the machine because of time-critical calculations. The Allocation Committee will have the power to adjust quotas over the years to reflect evolution in personnel (new faculty hires, changing postdocs, evolving research priorities, etc).

3 Data Management

Two types of data collections will need to be managed for the proposed facility. First is the data generated on the machine itself as result of ongoing local research. Second is the data generated remotely that needs to be archived locally and made publicly available. Based upon the characteristics of simulations we expect to run, the first class of data collections will be medium sized (a few terabytes per simulation set per group per year). The second type of data collection results from full production runs on the largest supercomputers in the US and is at least an order of magnitude larger.

For local simulations, fast, reliable storage is required for caching and staging the data during the run. Archival data collections, due to the relative infrequency of access, do not need to reside on such robust or fast devices and can thus use cheaper media. Our proposed solution is thus a hybrid system that provides rapid reliable access for a small subset of the data, but slower access for larger archived data sets, up to a total of approximately 1 PB. After local simulations are completed, data can be moved from the the fast cache storage to the slower archival storage. This optimal configuration will be achieved by the purchase of additional storage devices in the second year.

For archival storage, we will use SATA drives, instead of the much more expensive (but slightly more reliable) SAS drives. Each drive will have a capacity of 2 TB or higher. These SATA drives will be hosted in disk enclosures and connected to 8 file servers with SAS interfaces. We plan to recycle the 4 file servers and disk enclosures (but will replace the drives) from the current *Pleiades* cluster; thus we will only need to purchase four more file servers. The 1 PB storage will be served to the whole cluster using a fast parallel file system. We are exploring using Lustre, which is the *de facto* parallel file system, used by many of the top 500 machines.

In the second year of the grant we will add 10 TB of solid state drives (SSD) for caching and scratch space. This will greatly increase the performance of the overall storage system at modest cost. The caching/scratching storage will provide an expected IO speed of 10 GB/s (more than 10 times that of the main storage); and its total capacity is more than sufficient to hold all the output of running simulations for a short period of time (10–30 minutes) before the data are moved to the more spacious main storage.

The storage system will be directly linked to UCSC's "dark fiber network", which connects to the CalREN-HPR network tier at a bandwidth of 10 Gb/s. The dark fiber will allow the researchers to move large amount of data between the new cluster and supercomputers at national centers. It will also enable the proposed machine to serve as a repository for astrophysics legacy simulations, making these accessible to the community at large.

Within the budget of this proposal, extensive redundant backup systems (such as robotic tape storage) are not affordable. Our hybrid system of solid state storage for the necessarily fast and robust caching and staging process, and disk storage under RAID6 for the archival storage mitigates this issue. The greatly increased reliability of RAID6 will reduce the need for frequent backups, thus reducing cost. Moreover, we will be able to reduce the need for "hot spares" and fast rebuild by being able to survive two simultaneous disk failures, again reducing management cost. Since the archival storage will not be written frequently, the slightly lower write performance of RAID6 as compared to RAID5 will not be a major concern.

While the long-term storage systems will all be RAID6, and therefore protect against many device failures, we again intend to operate in a redundant "dual mode" with regards to backup of the permanently valuable data. The most valuable data are the large production collections and the facility here is intended as both a backup and public accessible system for this data. The data will also be kept at the production facility when possible, so that redundant copies exist. Our facility may outlive the remote entity however, at which point further backups become an issue. For the locally-generated data, no redundant backups (apart from the RAID) will exist. We anticipate that this is not a catastrophic issue since these simulations, though permanently valuable, are, by the nature of the machine, mid-range calculations, and therefore reasonably reproducible by recalculation (208), based on work done by some of the computer scientists in our team.

We expect that the archival repository will remain available for at least 3–4 years after the end of the proposal funding period. While we cannot guarantee funding for additional hardware during the post-grant period, the space, power, and network connectivity allocated to the system will remain available. Thus, we can continue to run the storage system using existing hardware; our experience with *Pleiades* suggests that the system will remain viable for 3 or more years beyond the end of the grant period. In addition, we will seek additional funding to extend the lifetime of the archive. Since we are doing work on archival storage systems as well as astrophysics, we may be able to leverage additional funding in either of these areas to build a longer-term storage solution. As noted above, we are also storing many important results at high-performance computing centers; thus, we do not expect to lose all data even after this system becomes too old to be useful.

4 References Cited

4.1 Publications Enabled by Pleiades since 2007

[1] Abdo, A. A., et al. (Fermi Collaboration) 2009, "Detection of 16 Gamma-Ray Pulsars Through Blind Frequency Searches Using the Fermi LAT", Science, 325, 840

[2] Anderson, B., Kuhlen, M., Diemand, J., Johnson, R., & Madau, P. 2010, "Fermi-LAT sensitivity to dark matter annihilation in Via Lactea II substructure", ApJ, 718, 899

[3] Aspden, A. J., Bell, J. B., Dong, S., & Woosley, S. E. 2011, "Burning Thermals in Type Ia Supernovae", ApJ, 738, 94

[4] Aspden, A. J., Bell, J. B., & Woosley, S. E. 2011, "Turbulent Oxygen Flames in Type Ia Supernovae", ApJ, 730, 144

[5] Asphaug, E. 2009, 'Growth and Evolution of Asteroids", Ann. Rev. Earth Plan. Sci., 37, 413

[6] Asphaug, E. 2010, "Invited Review: Similar sized collisions and the diversity of planets", Chemie der Erde, 70, 199

[7] Asphaug, E., Jutzi, M., & Movshovitz, N. 2011. "Chondrule formation during planetesimal accretion", Earth Planet. Sci. Lett., doi:10.1016/j.epsl.2011.06.007

[8] Baruteau, C., & Lin, D. N. C. 2010, "Protoplanetary Migration in Turbulent Isothermal Disks", ApJ, 709, 759

[9] Baruteau, C., Cuadra, J., & Lin, D. N. C. 2011, "Binaries Migrating in a Gaseous Disk: Where are the Galactic Center Binaries?", ApJ, 726, 28

[10] Batygin, K., & Laughlin, G. 2008, "On the Dynamical Stability of the Solar System", ApJ, 683, 1207

[11] Brummell, N.H., Tobias, S.M. & Cattaneo, F. 2010, "Dynamo efficiency in compressible dynamos with and without penetration", GAFD, 104, 565

[12] Brummell, N.H., Tobias, S.M., Thomas, J.H. & Weiss, N.O. 2008, "Flux pumping and magnetic fields in the outer penumbra of a sunspot", A&A, 686, 1454

[13] Busha, M. T., Marshall, P. J., Wechsler, R. H., Klypin, A. & Primack, J. 2011, "The Mass Distribution and Assembly of the Milky Way from the Properties of the Magellanic Clouds", ApJ, 743, 40

[14] Busha, M. T., Wechsler, R. H., Behroozi, P. S., Gerke, B. F., Klypin, A. A. & Primack, J. R., 2011, "Statistics of Satellite Galaxies Around Milky Way-Like Hosts", ApJ, 743, 117

[15] Chambers, L.S., Cuzzi, J.N., Asphaug, E., Colwell, J., & Sugita, S. 2008, "Hydrodynamical and radiative transfer modeling of meteoroid impacts into Saturn's rings", Icarus, 194, 623

[16] Chen, X., Madau, P., Sesana, A., & Liu, F. K. 2009, "Enhanced tidal disruption rates from massive black hole binaries", ApJ, 697, L149

[17] Chen, X., Sesana, A., Madau, P., & Liu, F. K. 2011, "Tidal stellar disruptions by massive black hole pairs: II. Decaying binaries", ApJ, 729, 13

[18] Covington, M., Dekel, A., Cox, T. J., Jonsson, P. & Primack, J. R. 2008, "Predicting the properties of the remnants of dissipative galaxy mergers", MNRAS, 384, 94

[19] Covington, M. D., Kassin, S. A., Dutton, A. A., Weiner, B. J., Cox, T. J., Jonsson, P., Primack, J. R., Faber, Sandra M., Koo, & D. C. 2010, "Evolution of the Stellar Mass Tully-Fisher Relation in Disk Galaxy Merger Simulations", ApJ, 710, 279

[20] Covington, M. D., Primack, J. R., Porter, L. A., Croton, D. J., Somerville, R. S. & Dekel, A. 2011, "The role of dissipation in the scaling relations of cosmological merger remnants", MNRAS, 415, 3135

[21] Cox, T. J., Jonsson, P., Somerville, R. S., Primack, J. R., & Dekel, A. 2008, "The effect of galaxy mass ratio on merger-driven starbursts", MNRAS, 384, 386

[22] Cuesta, A. J., Jeltema, T. E., Zandanel, F., Profumo, S., Prada, F., Yepes, G., Klypin, A., Hoffman, Y., Gottlber, S., Primack, J. R., Snchez-Conde, M. A., & Pfrommer, C. 2011, "Dark Matter Decay and Annihilation in the Local Universe: Clues from Fermi", ApJL, 726, L6 i

[23] Cunningham, A. J., Klein, R. I., McKee, C. F., & Krumholz, M. R. 2011, "Radiation-Hydrodynamic Simulations of Massive Star Formation with Protostellar Outflows", ApJ, 740, 107

[24] da Silva, R. L., Fumagalli, M., & Krumholz, M. R. 2012, "SLUGStochastically Lighting Up Galaxies. I. Methods and Validating Tests", ApJ, 745, 145

[25] Dan, M., Rosswog, S., Guillochon, J., & Ramirez-Ruiz, E. 2011, "Prelude to A Double Degenerate Merger: The Onset of Mass Transfer and Its Impact on Gravitational Waves and Surface Detonations", ApJ, 737, 89

[26] Diemand, J., Kuhlen, M., & Madau, P. 2007, "Dark matter substructure and γ-ray annihilation in the Milky Way halo", ApJ, 657, 262

[27] Diemand, J., Kuhlen, M., & Madau, P. 2007, "Formation and evolution of galaxy dark matter halos and their substructure", ApJ, 667, 859

[28] Diemand, J., Kuhlen, M., Madau, P., Zemp, M., Moore, B., Potter, D., & Stadel, J. 2008, "Clumps and streams in the local dark matter distribution", Nature, 454, 735

[29] Dobbs-Dixon, I., Cumming, A., & Lin, D. N. C. 2010, "Radiative Hydrodynamic Simulations of HD209458b: Temporal Variability", ApJ, 710, 1395

[30] Dobbs-Dixon, I., Li, S., & Lin, D. N. C. 2007, "Tidal barrier and the asymptotic mass of proto-gas giant planets", ApJ, 660, 791

[31] Dobbs-Dixon, I., & Lin, D. N. C. 2008, "Atmospheric Dynamics of Short-Period Extrasolar Gas Giant Planets. I. Dependence of Nightside Temperature on Opacity, ApJ, 673, 513

[32] Dominguez, A., Primack, J. R., et al. 2011, "Extragalactic background light inferred from AEGIS galaxy-SED-type fractions", MNRAS, 410, 2556

[33] Dukes, D., & Krumholz, M. R. 2012, "Was the Sun born in a massive cluster?", ApJ, submitted (arXiv:111.3693)

[34] Durda, D.D., Bottke, W.F., Nesvorny, D., Enke, B.L., Merline, W.J., Asphaug, E., & Richardson, D.C. 2007, "Size frequency distributions of fragments from SPH/N-body simulations of asteroid impacts: Comparison with observed asteroid families", Icarus, 186, 498

[35] Durda, D.R., Movshovitz, N., Richardson, D.C., Asphaug, E., Morgan, A., Rawlings, A.R. and Vest, C. 2011, "Experimental determination of the coefficient of restitution for meter-scale granite spheres", Icarus, 211, 849

[36] Evonuk, M. & Glatzmaier, G.A. 2007, "The effects of rotation on deep convection in giant planets with small cores", Planet. Space Sci. 55, 407

[37] Fellhauer, M., & Lin, D. N. C. 2007, "The influence of mass-loss from a star cluster on its dynamical friction - I. Clusters without internal evolution", MNRAS, 375, 604

[38] Flores, R. A., Allgood, B., Kravtsov, A. V., Primack, J. R., Buote, D. A., & Bullock, J. S. 2007, "The shape of galaxy cluster dark matter haloes: systematics of its imprint on cluster gas and comparison to observations", MNRAS, 377, 883

[39] Fortney, J. J., Shabram, M., Showman, A. P., Lian, Y., Freedman, R. S., Marley, M. S., & Lewis, N. K., 2010, "Transmission Spectra of Three-Dimensional Hot Jupiter Model Atmospheres", ApJ, 709, 1396

[40] Fumagalli, M., da Silva, R. L., & Krumholz, M. R. 2011, "Stochastic Star Formation and a (Nearly) Uniform Stellar Initial Mass Function", ApJL, 741, L26

[41] Fumagalli, M., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D. & Primack, J. R. 2011, "Absorption line systems in simulated galaxies fed by cold streams," MNRAS, 418, 1796

[42] Garaud, P. 2007, "Constraints on angular momentum transport in the Sun from simulations of the tachocline", Astron. Nachr., 328, 1146

[43] Garaud, P., & Acevedo-Arreguin, L. A. 2009, "On the Penetration of Meridional Circulation Below the Solar Convection Zone. II. Models with Convection Zone, the Taylor-Proudman Constraint, and Applications to Other Stars", ApJ, 704, 1

[44] Garaud, P., & Bodenheimer, P. 2010, "Gyroscopic Pumping of Large-scale Flows in Stellar Interiors and Application to Lithium-dip Stars", ApJ, 719, 313

[45] Garaud, P., & Brummell, N. H. 2008, "On the penetration of meridional circulation below the solar convection zone", ApJ, 674, 498

[46] Garaud, P., & Garaud, J.D. 2008, "Dynamics of the solar tachocline - II. The stratified case", MNRAS, 391, 1239

[47] Garaud, P., & Guervilly, C. 2009, "The Rotation Rate of the Solar Radiative Zone", ApJ, 695, 799

[48] Garaud, P., & Lin, D. N. C. 2007, "The effect of internal dissipation and surface irradiation on the structure of disks and the location of the snow line around sun-like stars", ApJ, 654, 606

[49] Garaud, P. , Ogilvie, G.I., Miller, N., & Stellmach, S. 2010, "A model of the entropy flux and Reynolds stress in turbulent convection", MNRAS, 407, 2451

[50] Gendelev, L., & Krumholz, M. R. 2012, "Evolution of blister-type H II regions in a magnetized medium", ApJ, 745, 158

[51] Gilmore, R. C., Madau, P., Primack, J. R., Somerville, R. S., & Haardt, F. 2009, "GeV gamma-ray attenuation and the high-redshift UV background," MNRAS, 399, 1694

[52] Gilmore, R. C., Prada, F., & Primack, J. R. 2010, "Modelling gamma-ray burst observations by Fermi and MAGIC including attenuation due to diffuse background light", MNRAS, 402, 565

[53] Gilmore, R. C., Primack, J. R., Bouvier, A., & Otte, A. N. 2011, Catching GRBs with atmospheric Cherenkov telescopes", AIPC, 1358, 213

[54] Gilmore, R. C., Somerville, R. S., Primack, J. R., & Dominguez, A. 2012, "Semi-analytic modeling of the EBL and consequences for extragalactic jgamma-ray spectra", MNRAS, in press (arXiv:1104.0671)

[55] Goerdt, T., Dekel, A., Sternberg, A., Ceverino, D., Teyssier, R. & Primack, J. R., 2010, "Gravity-driven Lyman-alpha blobs from cold streams into galaxies", MNRAS, 407, 613

[56] Glatzmaier, G.A. 2008, "A note on 'Constraints on deep-seated zonal winds inside Jupiter and Saturn", Icarus, 196, 665

[57] Glatzmaier, G.A. & Coe, R.S. 2007, "Magnetic polarity reversals in the core", in Treatise on Geophysics, Volume 8, Core Dynamics, Chp. 9, ed. G. Schubert (Elsevier), pp. 283

[58] Glatzmaier, G.A., Evonuk, M. & Rogers, T.M. 2009, "Differential rotation in giant planets maintained by density-stratified turbulent convection", GAFD, 103, 31

[59] Gritschneder, M., Lin, D. N. C., Murray, S. D., Yin, Q.-Z., & Gong, M.-N. 2012, "The Supernova Triggered Formation and Enrichment of Our Solar System", ApJ, 745, 22

[60] Guedes, J., Callegari, S., Madau, P., & Mayer, L. 2011, "Forming realistic late-type spirals in a ΛCDM universe: The Eris simulation", ApJ, 742, 76

[61] Guedes, J., Madau, P., Kuhlen, M., Diemand, J., & Zemp, M. 2009, "Simulations of recoiling massive black holes in the Via Lactea halo", ApJ, 702, 890

[62] Guedes, J., Madau, P., Mayer, L., & Callegari, S. 2011, "Recoiling massive black holes in gas-rich galaxy merger remnants", ApJ, 729, 125

[63] Guedes, J. M., Rivera, E. J., Davis, E., Laughlin, G., Quintana, E., Fischer, D. A. 2008, "Formation and Detectability of Terrestrial Planets around α Centauri B", ApJ, 679, 1582

[64] Guillochon, J., Ramirez-Ruiz, E., & Lin, D. 2011, "Consequences of the Ejection and Disruption of Giant Planets", ApJ, 732, 74

[65] Guillochon, J., Ramirez-Ruiz, E., Rosswog, S., & Kasen, D. 2009, "Three-dimensional Simulations of Tidally Disrupted Solar-type Stars and the Observational Signatures of Shock Breakout", ApJ, 705, 844

[66] Guillochon, J., Dan, M., Ramirez-Ruiz, E., & Rosswog, S. 2010, "Surface Detonations in Double Degenerate Binary Systems Triggered by Accretion Stream Instabilities", ApJ, 709, L64

[67] Hennebelle, P., Commerçon, B., Joos, M., Klessen, R. S., Krumholz, M. R., Tan, J. C., & Teyssier, R. 2011, "Collapse, Outflows, and Fragmentation of Massive, Turbulent, and Magnetized Prestellar Barotropic Cores", A&A, 528, 72

[68] Ida, S., & Lin, D. N. C. 2008, 'Toward a Deterministic Model of Planetary Formation. IV. Effects of Type I Migration", ApJ, 673, 487

[69] Ida, S., & Lin, D. N. C. 2008, "Toward a Deterministic Model of Planetary Formation. V.. Accumulation Near the Ice Line and Super-Earths", ApJ, 685, 584

[70] Ida, S., & Lin, D. N. C. 2010, "Toward a Deterministic Model of Planetary Formation. VI. Dynamical Interaction and Coagulation of Multiple Rocky Embryos and Super-Earth Systems around Solar-type Stars", ApJ, 719, 810

[71] Joggerst, C. C., Almgren, A., Bell, J., Heger, A., Whalen, D., & Woosley, S. E. 2010, "The Nucleosynthetic Imprint of 15-40 M_{sun} Primordial Supernovae on Metal-Poor Stars", ApJ, 709, 11

[72] Joggerst, C. C., Almgren, A., & Woosley, S. E. 2010, "Three-dimensional Simulations of Rayleigh-Taylor Mixing in Core-collapse Supernovae with Castro", ApJ, 723, 353

[73] Joggerst, C. C., Woosley, S. E., & Heger, A. 2009, "Mixing in Zero- and Solar-Metallicity Supernovae", ApJ, 693, 1780

[74] Jones, C.A., Boronski, P., Brun, A.S., Glatzmaier, G.A., Gastine, T., Miesch, M.S., & Wicht, J. 2011, "Anelastic convection-driven dynamo benchmarks", Icarus, 216, 120

[75] Jonsson, P., Cox, T. J., Primack, J. R. & Somerville, R. S., 2006. Simulations of Dust in Interacting Galaxies. I. Dust Attenuation. ApJ, 637, 255.

[76] Jonsson, P. & Primack, J. R., 2010, "Accelerating dust temperature calculations with graphics-processing units", New Astronomy, 15, 509

[77] Jutzi, M. & Asphaug, E. 2011. "Forming the lunar farside highlands by the accretion of a companion moon", Nature, 476, 69-72

[78] Jutzi, M. & Asphaug, E. 2011. "Mega-ejecta on asteroid Vesta", Geophys. Res. Lett., 38, L01102

[79] Kasen, D., & Ramirez-Ruiz, E. 2010, "Optical Transients from the Unbound Debris of Tidal Disruption", ApJ, 714, 155

[80] Kasen, D., Roepke, F. K., & Woosley, S. E. 2009, "The diversity of type Ia supernovae from broken symmetries", Nature, 460, 869

[81] Kasen, D., & Woosley, S. E. 2009, "Type II Supernovae: Model Light Curves and Standard Candle Relationships", ApJ, 703, 2205

[82] Kasen, D., Thomas, R. C., Roepke, F., & Woosley, S. E. 2008, "Multidimensional radiative transfer calculations of the light curves and spectra of Type Ia supernovae", Journal of Physics Conference Series, 125, 012007

[83] Kasen, D., 2010, "Seeing the Collision of a Supernova with Its Companion Star", ApJ, 708, 1025

[84] Kasen, D., Woosley, S.E., & Heger, A., 2011, "Light Curves of Pair-Instability Supernovae", ApJ, 734, 102

[85] Ke, T. T., Huang, H., & Lin, D. N. C. 2012, "Rapid Mid-infrared Variability in Protostellar Disks", ApJ, 745, 60

[86] Kelley, L. Z., Ramirez-Ruiz, E., Zemp, M., Diemand, J., & Mandel, I. 2010, "The Distribution of Coalescing Compact Binaries in the Local

Universe: Prospects for Gravitational-wave Observations", ApJ, 725, L91

[87] Klypin, A. A., Trujillo-Gomez, S. & Primack, J. 2011, "Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation", ApJ, 740, 102

[88] Korycansky, D.G., Plesko, C.S., Jutzi, M., Asphaug, E. & Colaprete, A. 2009, "Predictions for the LCROSS Mission", Meteoritics and Plan. Sci., 44, 603

[89] Kratter, K. M., Matzner, C. D., Krumholz, M. R., & Klein, R. I. 2010, "On the Role of Disks in the Formation of Stellar Systems: A Numerical Parameter Study of Rapid Accretion", ApJ, 708, 1585

[90] Kretke, K. A., & Lin, D. N. C. 2007, "Grain retention and formation of planetesimals near the snow line in MRI-driven turbulent protoplanetary disks", ApJ, 664, 55

[91] Kretke, K. A., & Lin, D. N. C. 2010, "Structure of Magnetorotational Instability Active Protoplanetary Disks", ApJ, 721, 1575

[92] Kretke, K. A., Lin, D. N. C., Garaud, P., & Turner, N. J. 2009, "Assembling the Building Blocks of Giant Planets Around Intermediate-Mass Stars", ApJ, 690, 407

[93] Krumholz, M. R., Cunningham, A. J., Klein, R. I., & McKee, C. F. 2010, "Radiation Feedback, Fragmentation, and the Environmental Dependence of the Initial Mass Function", ApJ, 713, 1120

[94] Krumholz, M. R., & Gnedin, N. Y. 2011, "A Comparison of Methods for Determining the Molecular Content of Model Galaxies", ApJ, 729, 36

[95] Krumholz, M. R., Klein, R. I., & McKee, C. F. 2011, "Radiation-Hydrodynamic Simulations of the Formation of Massive Star Clusters I. Implications for the Origin of the Initial Mass Function", ApJ, 740, 74

[96] Krumholz, M. R., Klein, R. I., McKee, C. F., Offner, S. S. R., & Cunningham, A. J. 2009, "The Formation of Massive Star Systems by Accretion", Science, 323, 754

[97] Krumholz, M. R., Stone, J. M & Gardiner, T. A. 2007, "Magnetohydronamic Evolution of HII Regions: Simulation Methodology, Convergence Tests, and Uniform Media", ApJ, 671, 518

[98] Krumholz, M. R., Klein, R. I., McKee, C. F., & Bolstad, J. 2007, "Equations and Algorithms for Mixed-Frame Flux Limited Diffusion Radiation Hydrodynamics", ApJ, 667, 626

[99] Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, "Molecular Line Emission from Massive Protostellar Disks: Predictions for ALMA and the EVLA", ApJ, 665, 478

[100] Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, "Simulations of Collapse and Fragmentation in Massive Protostellar Cores", ApJ, 656, 959

[101] Kuhlen, M., Diemand, J., & Madau, P. 2007, "The shapes, orientation, and alignment of galactic dark matter subhalos", ApJ, 671, 1135

[102] Kuhlen, M., Diemand, J., & Madau, P. 2008, "The dark matter annihilation signal from galactic substructure: predictions for GLAST", ApJ, 686, 262

[103] Kuhlen, M., Diemand, J., Madau, P., & Zemp, M. 2008, "The Via Lactea INCITE simulation: galactic dark matter substructure at high resolution", JPhCS, 125, 012008

[104] Kuhlen, M., Krumholz, M., Madau, P., Smith, B. D., & Wise, J. H. 2012, "Dwarf galaxy formation with H_2-regulated star formation", ApJ, submitted (arXiv:1105.2376)

[105] Kuhlen, M., Madau, P., & Silk, J. 2009, "Exploring the dark sector with Milky Way substructure", Science, 325, 970

[106] Kuhlen, M., Weiner, N., Diemand, J., Madau, P., Moore, B., Potter, D., Stadel, J., & Zemp, M. 2010, "Dark matter direct detection with non-Maxwellian velocity structure", JCAP, 02, 030

[107] Laine, Randy O., & Lin, Douglas N. C. 2012, "Interaction of Close-in Planets with the Magnetosphere of Their Host Stars. II. Super-Earths as Unipolar Inductors and Their Orbital Evolution", ApJ, 745, 2

[108] Laine, Randy O., Lin, Douglas N. C., & Dong, Shawfeng 2008, "Interaction of Close-in Planets with the Magnetosphere of Their Host Stars. I. Diffusion, Ohmic Dissipation of Time-dependent Field, Planetary Inflation, and Mass Loss", ApJ, 685, 521

[109] Langton, J., Laughlin, G. 2008, "Hydrodynamic Simulations of Unevenly Irradiated Jovian Planets", ApJ, 674, 1106

[110] Langton, J., Laughlin, G. 2008, "Persistent circumpolar vortices on the extrasolar giant planet HD 37605 b", A&A, 483, L25

[111] Laughlin, G., Deming, D., Langton, J., Kasen, D., Vogt, S., Butler, P., Rivera, E., Meschiari, S. 2009, "Rapid heating of the atmosphere of an extrasolar planet", Nature 457, 562

[112] Lee, W. H., & Ramirez-Ruiz, E. 2007, "The progenitors of short gamma-ray bursts", New Journal of Physics, 9, 17

[113] Lee, W. H., Ramirez-Ruiz, E., & van de Ven, G. 2010, "Short Gamma-ray Bursts from Dynamically Assembled Compact Binaries in Globular Clusters: Pathways, Rates, Hydrodynamics, and Cosmological Setting", ApJ, 720, 953

[114] Lee, W. H., Ramirez-Ruiz, E., & López-Cámara, D. 2009, "Phase Transitions and He-Synthesis-Driven Winds in Neutrino Cooled Accretion Disks: Prospects for Late Flares in Short Gamma-Ray Bursts", ApJ, 699, L93

[115] Lewis, N. K., Showman, A. P., Fortney, J. J., Marley, M. S., Freedman, R. S., & Lodders, K., 2010, "Atmospheric Circulation of Eccentric Hot Neptune GJ436b", ApJ, 720, 344

[116] Li, H., Lubow, S. H., Li, S., & Lin, D. N. C. 2009, "Type I Planet Migration in Nearly Laminar Disks", ApJ, 690, L52

[117] Li, S. L., Agnor, C. B., & Lin, D. N. C. 2010, "Embryo Impacts and Gas Giant Mergers. I. Dichotomy of Jupiter and Saturn's Core Mass", ApJ, 720, 1161

[118] Lin, M.-K., Krumholz, M. R., & Kratter, K. M. 2011, "Spin Down of Protostars Through Gravitational Torques", MNRAS, 416, 580

[119] Lin, D. N. C., & Murray, S. D. 2007, "Gas accretion by globular clusters and nucleated dwarf galaxies and the formation of the arches and quintuplet clusters", ApJ, 661, 779

[120] Lopez-Camara, D., Lee, W. H., & Ramirez-Ruiz, E. 2009, "Gamma-Ray Burst Production and Supernova Signatures in Slowly Rotating Collapsars", ApJ, 692, 804

[121] Lopez-Camara, D., Lee, W. H., & Ramirez-Ruiz, E. 2010, "Critical Angular Momentum Distributions in Collapsars: Quiescent Periods from Accretion State Transitions in Long Gamma-ray Bursts", ApJ, 716, 1308

[122] Lotz, J. M., Jonsson, P., Cox, T. J., Croton, D., Primack, J. R., Somerville, R. S. & Stewart, K. 2011, "The Major and Minor Galaxy Merger Rates at $z \leq 1.5$", ApJ, 742, 103

[123] Lotz, J. M., Jonsson, P., Cox, T. J. & Primack, J. R., 2008, "Galaxy merger morphologies and time-scales from simulations of equal-mass gas-rich disc mergers", MNRAS, 391, 1137

[124] Lotz, J. M., Jonsson, P., Cox, T. J. & Primack, J. R., 2010, "The effect of gas fraction on the morphology and time-scales of disc galaxy mergers", MNRAS, 404, 590

[125] Lotz, J. M., Jonsson, P., Cox, T. J. & Primack, J. R., 2010, "The effect of mass ratio on the morphology and time-scales of disc galaxy mergers", MNRAS, 404, 575

[126] Lu, Y., Yu, Q., & Lin, D. C. N. 2007, "Hypervelocity binary stars: smoking gun of massive binary black holes", ApJ, 666, 89

[127] Madau, P., Diemand, J., & Kuhlen, M. 2008, "Dark matter subhalos and the dwarf satellites of the Milky Way", ApJ, 679, 1260

[128] Madau, P., Kuhlen, M., Diemand, J., Moore, B., Zemp, M., Potter, D., & Stadel, J. 2008, "Fossil remnants of reionization in the halo of the Milky Way", ApJ, 689, L41

[129] Marinova, M.M., Aharonson, O., & Asphaug, E. 2008, "Mega-impact formation of the Mars hemispheric dichotomy", Nature, 453, 1216

[130] Marinova, M.M., Aharonson, O., & Asphaug, E. 2011, "Geophysical consequences of planetary-scale impacts into a Mars-like planet", Icarus, 211, 960

[131] Mayer, L., Kazantzidis, S., Madau, P., Colpi, M., Quinn, T., & Wadsley, J. 2007, "Rapid formation of supermassive black hole binaries in galaxy mergers with gas", Science, 316, 1874

[132] Meschiari, S., Laughlin, G. 2010, "Systemic: A Testbed for Characterizing the Detection of Extrasolar Planets. II. Numerical Approaches to the Transit Timing Inverse Problem", ApJ, 718, 543

[133] Muñoz, J. A., Madau, P., Loeb, A., & Diemand, J. 2009, "Probing reionization with Milky Way satellites", MNRAS, 400, 1593

[134] Myers, A. T., Krumholz, M. R., Klein, R. I., & McKee, C. F. 2011, "Metallicity and the Universality of the IMF", ApJ, 735, 49

[135] Naiman, J. P., Ramirez-Ruiz, E., & Lin, D. N. C. 2011, "External Mass Accumulation onto Core Potentials: Implications for Star Clusters, Galaxies, and Galaxy Clusters", ApJ, 735, 25

[136] Naiman, J. P., Ramirez-Ruiz, E., & Lin, D. N. C. 2009, "Gas Accretion by Star Clusters and the Formation of Ultraluminous X-Ray Sources from Cusps of Compact Remnants", ApJ, 705, L153

[137] Narayanan, D., Krumholz, M. R., Ostriker, E. C., & Hernquist, L. 2011, "The CO-H_2 Conversion Factor in Disc Galaxies and Mergers", MNRAS, 418, 664

[138] Nimmo, F., Hart, S.D., Korycansky, D.G., & Agnor, C.B. 2008, "Implications of an impact origin for the Martian hemispheric dichotomy", Nature, 453, 1220

[139] Novak, G., Cox, T. J., Primack, J. R., & Dekel, A. 2006, "Shapes of Stellar Systems and Dark Halos from Simulations of Galaxy Major Mergers", ApJ, 646, L9

[140] Offner, S. S. R., Hansen, C., & Krumholz, M. R. 2009, "Stellar Kinematics of Young Clusters in Turbulent Hydrodynamic Simulations", ApJL, 704, L124

[141] Offner, S. S. R., Klein, R. I., McKee, C. F., & Krumholz, M. R. 2009, "The Effects of Radiative Transfer on Low-Mass Star Formation", ApJ, 703, 131

[142] Offner, S. S. R., Kratter, K. M., Matzner, C. D., Krumholz, M. R., & Klein, R. I. 2010, "The Formation of Low-Mass Binary Star Systems Via Turbulent Fragmentation", ApJ, 725, 1485

[143] Offner, S. S. R., & Krumholz, M. R. 2009, "The Shapes of Molecular Cloud Cores in Simulations and Observation", ApJ, 693, 914

[144] Offner, S. S. R., Krumholz, M. R., Klein, R. I., & McKee, C. F. 2008, "The Dynamics of Molecular Cloud Cores in Driven and Undriven Turbulence Environments", AJ, 136, 404

[145] Ogilvie, G., & Lin, D. N. C, 2007, "Tidal dissipation in rotating solar-type stars", ApJ, 661, 1180

[146] Olson, P.L., Coe, R.S., Driscall, P.E., Glatzmaier, G.A. & Roberts, P.H. 2010, "Geodynamo reversal frequency and heterogeneous core-mantle boundary heat flow", Phys. Earth Planet. Inter., 180, 66

[147] Olson, P.L., Glatzmaier, G.A., & Coe, R.S. 2011, "Complex polarity reversals in a geodynamo model", Earth Planet. Sci. Lett., 304, 168

[148] Ong, L., Asphaug, E., Korycansky, D., & Coker, R.F. 2010, "Volatile retention from cometary impacts on the Moon", Icarus, 207, 578

[149] Pierazzo, E., Artemieva, N., Asphaug, E., Baldwin, E.C., Cazamias, J., Coker, R., Collins, G.S., Crawford, D.A., Davison, T., Elbeshausen, D., Holsapple, K.A., Housen, K.R., Korycansky,

D.G., & Wnnemann , K. 2008, “Validation of numerical codes for impact and explosion cratering: Impacts on strengthless and metal targets”, Meteoritics and Plan. Sci., 43, 1917

[150] Pierce, C. M., Lotz, J. M., Primack, J. R., et al. 2010, “The effects of an active galactic nucleus on host galaxy colour and morphology measurements”, MNRAS, 405, 718

[151] Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E. & Primack, J. 2011, “Halo concentrations in the standard LCDM cosmology”, MNRAS, submitted (arXiv1106.5587)

[152] Ramirez-Ruiz, E., & MacFadyen, A. I. 2010, “The Hydrodynamics of Gamma-ray Burst Remnants”, ApJ, 716, 1028

[153] Ramirez-Ruiz, E., Nishikawa, K.-I., & Hededal, C. B. 2007, “$e^{+/-}$ Pair Loading and the Origin of the Upstream Magnetic Field in GRB Shocks”, ApJ, 671, 1877

[154] Ramirez-Ruiz, E., & Rosswog, S. 2009, “The Star Ingesting Luminosity of Intermediate-Mass Black Holes in Globular Clusters”, ApJ, 697, L77

[155] Rashkov, V., Madau, P., Kuhlen, M., & Diemand, J. 2012, “On the assembly of the Milky Way dwarf satellites and their common mass scale”, ApJ, 745, 142

[156] Riebe, K., Partl, A. M., Enke, H., Forero-Romero, J., Gottloeber, S., Klypin, A., Lemson, G., Prada, F., Primack, J. R., Steinmetz, M. & Turchaninov, V. 2011, “The MultiDark Database: Release of the Bolshoi and MultiDark Cosmological Simulations”, New Astronomy, submitted (arXiv:1109.0003)

[157] Roberts, P.H., Glatzmaier, G.A. & Clune, T.L. 2010, “Numerical simulation of a spherical dynamo excited by a flow of von Karmon type”, GAFD, 104, 202

[158] Roberts, J.H., & Nimmo, F. 2008, “Near-surface heating on Enceladus and the south polar thermal anomaly”, Geophys. Res. Lett., 35, L09201

[159] Roberts, J.H., & Nimmo, F. 2008, “Tidal heating and the long-term stability of a subsurface ocean on Enceladus”, Icarus, 194, 675

[160] Roberts, L. F., Kasen, D., Lee, W. H., & Ramirez-Ruiz, E. 2011, “Electromagnetic Transients Powered by Nuclear Decay in the Tidal Tails of Coalescing Compact Binaries”, ApJL, 736, 21

[161] Roberts, L. F., Shen, G., Cirigliano, V., et al., 2011, “Proto-Neutron Star Cooling with Convection: The Effect of the Symmetry Energy”, Phys. Rev. Lettr., in press (arXiv:1112.0335)

[162] Robuchon, G., & Nimmo, F. 2011, “Thermal evolution of Pluto and implications for surface tectonics and a subsurface ocean”, Icarus, 216, 426

[163] Robuchon, G., Nimmo, F., Roberts, J., & Kirchoff, M. 2011, “Impact basin relaxation at Iapetus”, Icarus, 214, 82

[164] Robuchon, G., Zhang, K., & Nimmo, F. 2011, “Coupling satellite thermal and orbital evolution: Application to Enceladus and Dione”, EPSC-DPS2011-1178

[165] Rocha, Miguel, Jonsson, P., Primack, J. R. & Cox, T. J. 2008, “Dust attenuation in hydrodynamic simulations of spiral galaxies,” MNRAS, 383, 1281

[166] Rogers, T.M., MacGregor, K.B. & Glatzmaier, G.A. 2008, “Nonlinear dynamics of gravity wave driven flows in the solar radiative interior”, MNRAS, 387, 616

[167] Rosenblum, E. , Garaud, P. , Traxler A. L. & Stellmach, S. 2011, “Turbulent mixing and layer formation in double-diffusive convection: 3D numerical simulations and theory”, ApJ, 731, 66

[168] Rosswog, S., Kasen, D., Guillochon, J., & Ramirez-Ruiz, E. 2009, “Collisions of White Dwarfs as a New Progenitor Channel for Type Ia Supernovae”, ApJ, 705, L128

[169] Rosswog, S., Ramirez-Ruiz, E., Hix, W. R., & Dan, M. 2008, “Simulating black hole white

dwarf encounters", Computer Physics Communications, 179, 184

[170] Rosswog, S., Ramirez-Ruiz, E., & Hix, W. R. 2008, "Atypical Thermonuclear Supernovae from Tidally Crushed White Dwarfs", ApJ, 679, 1385

[171] Rosswog, S., Ramirez-Ruiz, E., & Hix, W. R. 2009, "Tidal Disruption and Ignition of White Dwarfs by Moderately Massive Black Holes", ApJ, 695, 404

[172] Shen, S., Madau, P., Aguirre, A., Guedes, J., Mayer, L., & Wadsley, J. 2011, "The origin of metals in the circum-galactic medium of massive galaxies at $z = 3$", ApJ, submitted (arXiv:1109.3713)

[172] Kuhlen, M., Krumholz, M., Madau, P., Smith, B. D., & Wise, J. H. 2012, "Dwarf galaxy formation with H_2-regulated star formation", ApJ, submitted (arXiv:1105.2376)

[173] Showman, A. P., Fortney, J. J., Lian, Y., Marley, M. S., Freedman, R. S., Knutson, H. A., & Charbonneau, D. 2009, "Atmospheric Circulation of Hot Jupiters: Coupled Radiative-Dynamical General Circulation Model Simulations of HD 189733b and HD 209458b", ApJ, 699, 564

[174] Showman, A. P., Cooper, C. S., Fortney, J. J., & Marley, M. S. 2008, "Atmospheric Circulation of Hot Jupiters: Three-dimensional Circulation Models of HD 209458b and HD 189733b with Simplified Forcing", ApJ, 682, 559

[175] Silvers, L.J., Vasil, G.M., Brummell, N.H., & Proctor, M.R.E. 2009, "Double-diffusive instabilities of a shear-generated magnetic layer", ApJ, 702, L14

[176] Socrates, A., Davis, S. W., & Ramirez-Ruiz, E. 2008, "The Eddington Limit in Cosmic Rays: An Explanation for the Observed Faintness of Starbursting Galaxies", ApJ, 687, 202

[177] Somerville, R. S., Gilmore, R. C., Primack, J. R., & Dominguez, A. 2012, "Galaxy Properties from the Ultra-violet to the Far-Infrared: Lambda-CDM models confront observations", MNRAS, in press (arXiv:1104.0671)

[178] Spurzem, R., Giersz, M., Heggie, D. C., & Lin, D. N. C. 2009, "Dynamics of Planetary Systems in Star Clusters", ApJ, 697, 458

[179] Stadel, J., Potter, D., Moore, B., Diemand, J., Kuhlen, M., Madau, P., Zemp, M., & Quilis, V. 2009, "Quantifying the heart of darkness with GHALO - a multi-billion particle simulation of our galactic halo", MNRAS, 398, L21

[180] Stanley, S. & Glatzmaier, G.A. 2009, "Dynamo models for planets other than Earth", Space Sci. Rev., doi:10.1007/s11214-009-9573-y.

[181] Stellmach, S., Traxler, A., Garaud, P., Brummell, N. & Radko, T. 2011, "Dynamics of fingering convection II: The formation of thermohaline staircases", JFM, 677, 554

[182] Suijs, M. P. L., Langer, N., Poelarends, A.-J., Yoon, S.-C., Heger, A., & Herwig, F. 2008, "White dwarf spins from low-mass stellar evolution models", A&A, 481, 87

[183] Tobias, S.M., Cattaneo, F. & Brummell, N.H. 2008, "Convective dynamos with penetration, rotation and shear", ApJ, 685, 596

[184] Tobias, S.M., Cattaneo, F. & Brummell, N.H. 2011, "On the generation of organized magnetic fields", ApJ, 728, 153

[185] Thommes, E., Nagasawa, M., & Lin, D. N. C. 2008, "Dynamical Shake-up of Planetary Systems. II. N-Body Simulations of Solar System Terrestrial Planet Formation Induced by Secular Resonance Sweeping", ApJ, 676, 728

[186] Traxler, A. L., Garaud, P. & Stellmach, S. 2011, "Numerically determined transport laws for fingering ("thermohaline") convection in astrophysics", ApJL, 728, L29

[187] Traxler, A., Stellmach, S., Garaud, P., Radko, T., & Brummell, N. 2011, "Dynamics of fingering convection I: Small-scale fluxes and large-scale instabilities", JFM, 677, 530

[188] Trujillo-Gomez, S., Klypin, A., Primack, J. & Romanowsky, A. J. 2011, "Galaxies in LCDM

with Halo Abundance Matching: Luminosity-Velocity Relation, Baryonic Mass-Velocity Relation, Velocity Function, and Clustering", ApJ, 742, 16

[189] Vasil, G. & Brummell, N.H. 2008, "Magnetic buoyancy instabilities of a shear-generated magnetic layer", ApJ, 686, 709

[190] Vasil, G. & Brummell, N.H. 2009, "Constraints on magnetic buoyancy instabilities of a shear-generated magnetic layer", ApJ, 690, 783

[191] Vasil, G., Brummell, N.H. & Julien, K. 2008a, "On the utility of fast transforms in parity-mixed PDEs: Part 1. Numerical techniques and analysis", JCP, 227, 7999

[192] Vasil, G., Brummell, N.H. & Julien, K. 2008b, "On the utility of fast transforms in parity-mixed PDEs: Part 2. Application to confined rotating convection", JCP, 227, 8017

[193] Watanabe, Sei-ichiro, & Lin, D. N. C. 2008, "Thermal Waves in Irradiated Protoplanetary Disks", ApJ, 672, 1183

[194] Woosley, S. E., & Kasen, D. 2011, "Sub-Chandrasekhar Mass Models for Supernovae", ApJ, 734, 38

[195] Woosley, S. E., Kasen, D., Blinnikov, S., & Sorokina, E. 2007, "Type Ia Supernova Light Curves", ApJ, 662, 487

[196] Woosley, S. E., & Zhang, W. 2007, "Models for GRBs and diverse transients", RSPTA, 365, 1129

[197] Yoon, S.-C., Langer, N., Cantiello, M., Woosley, S.E., & Glatzmaier, G. A. 2009, "Evolution of Progenitor Stars of Type Ibc Supernovae and Long Gamma-Ray Bursts", in Massive Stars as Cosmic Engines, IAU Symposium, 250, 231

[198] Yoon, S.-C., Woosley, S. E., & Langer, N. 2010, "Type Ib/c Supernovae in Binary Systems. I. Evolution and Properties of the Progenitor Stars", ApJ, 725, 940

[199] Yu, Q., & Madau, P. 2007, "Kinematics of hypervelocity stars in the triaxial halo of the Milky Way", MNRAS, 379, 1293

[200] Zemp, M., Diemand, J., Madau, P., Kuhlen, M., Moore, B., Potter, D., Stadel, J., & Widrow, L. 2009, "The graininess of dark matter halos", MNRAS, 394, 641

[201] Zemp, M., Moore, B., Stadel, J., Carollo, C. M., & Madau, P. 2008, "Multi-mass spherical structure models for N-body simulations", MNRAS, 386, 1543

[202] Zemp, M., Ramirez-Ruiz, E., & Diemand, J. 2009, "Halo Retention and Evolution of Coalescing Compact Binaries in Cosmological Simulations of Structure Formation: Implications for Short Gamma-ray Bursts", ApJ, 705, L186

[203] Zhang, Hui, Yuan, Chi, Lin, D. N. C., & Yen, David C. C. 2008, "On the Orbital Evolution of a Jovian Planet Embedded in a Self-Gravitating Disk", ApJ, 676, 739

[204] Zhang, K., & Nimmo, F. 2009, "Recent orbital evolution and the internal structures of Enceladus and Dione", Icarus, 204, 597

[205] Zhang, W., Woosley, S. E., & Heger, A. 2008, "Fallback and Black Hole Production in Massive Stars", ApJ, 679, 639

[206] Zhou, J.-L., & Lin, D. N. C. 2007, "Planetesimal accretion onto growing proto-gas giant planets", ApJ, 666, 447

[207] Zhou, J.-L., Lin, D. N. C., & Sun, Y.-S. 2007, "Post-oligarchic evolution of protoplanetary embryos and the stability of planetary systems", ApJ, 666, 423

4.2 Other References Cited

[208] Adams, I., Long, D. E., Miller, E.L., Pasupathy, S., and Storer, M. W., 2009, "Maximizing Efficiency By Trading Storage for Computation", Proceedings of the Workshop on Hot Topics in Cloud Computing (HotCloud '09)

[209] Adams, I. F., Miller, E. L., and Storer, M. W., 2010, "Examining Energy Use in Heterogeneous Archival Storage Systems", Proceedings of the 18th Annual Meeting of the IEEE International

Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2010)

[210] Adams, I. F., et al., 2012,"Analysis of Workload Behavior in Scientific and Historical Long-Term Data Repositories" ACM Transactions on Storage, 8, 2

[211] Buck, J. B., Watkins, N., LeFevre, J., Ioannidou, K., Maltzahn, C., Polyzotis, N., and Brandt, S., 2011, "SciHadoop: Array-based Query Processing in Hadoop, Proceedings of the 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC '11)

[212] Almgren, A. S., Beckner, V. E., Bell, J.B., et al., 2010, "CASTRO: A New Compressible Astrophysical Solver. I. Hydrodynamics and Self-gravity", ApJ, 715, 1221

[213] Governato, F., Brook, C., Mayer, L., Brooks, A., Rhee, G., Wadsley, J., Jonsson, P., Willman, B., Stinson, G., Quinn, T., & Madau, P. 2010, Bulgeless dwarf galaxies and dark matter cores from supernova-driven outflows", Nature, 463, 203

[214] Haardt, F., & Madau, P. 2012, "Radiative transfer in a clumpy universe: IV. New synthesis models of the cosmic UV/X-ray background", ApJ, in press (arXiv:1105.2039)

[215] Malone, C. M., Nonaka, A., Almgren, A. S., Bell, J. B., & Zingale, M. 2011, "Multidimensional Modeling of Type I X-ray Bursts. I. Two-dimensional Convection Prior to the Outburst of a Pure 4He Accretor", ApJ, 728, 118

[216] Jones, S. N., Strong, C. R., Parker-Wood, A., Holloway, A., Long, D. D. E., 2011, "Easing the Burdens of HPC File Management", in Proceedings of the 6th Petascale Data Storage Workshop (PDSW '11)

[217] Leung, A., Shao, M., Bisson, T., Pasupathy, S., and Miller, E. L. , "Spyglass: Fast, Scalable Metadata Search for Large-Scale Storage Systems", in Proceedings of the 7th USENIX Conference on File and Storage Technologies (FAST '09)

[218] Leung, A., Adams, I., and Miller, E. L., 2009, "Magellan: A Searchable Metadata Architecture for Large-Scale File Systems", University of California, Santa Cruz publication, UCSC-SSRC-09-07

[219] Nonaka, A., Almgren, A. S., Bell, J. B., Lijewski, M. J., Malone, C. M. & Zingale, M. 2010, "MAESTRO: An Adaptive Low Mach Number Hydrodynamics Algorithm for Stellar Flows", ApJS, 188, 358

[220] Parker-Wood, A., Strong, C., Miller, E. L., and Long, D. D. E., 2010, "Security Aware Partitioning for Efficient File System Search", 26th IEEE Symposium on Massive Storage Systems and Technologies: Research Track (MSST 2010)

[221] Rajendran, R., Miller, E. L., and Long, D. D. E., 2011, "Horus: Fine-Grained Encryption-Based Security for High Performance Petascale Storage", in Proceedings of the 6th Petascale Data Storage Workshop (PDSW '11)

[222] Wadsley, J. W., Stadel, J., & Quinn, T. 2004, "Gasoline: a flexible, parallel implementation of TreeSPH", New Astronomy, 9, 13

[223] Weil, S. A., Brandt, S. A., Miller, E. L., Long, D. D. E., Maltzahn, C., 2006, "Ceph: A Scalable, High-Performance Distributed File System", Proceedings of the 7th Symposium on Operating Systems Design and Implementation (OSDI '06)

[224] Wildani, A., Schwarz, T., Miller, E. L., and Long, D. D. E., 2009, "Protecting Against Rare Event Failures in Archival Systems", in Proceedings of the 17th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2009)

5 Budget Justification

The request from NSF – \$910,000 – is entirely for computer hardware. We are also budgeting \$130,000 from the matching funds for computer hardware. A detailed discussion of the hardware selected was given in § 1.4. Four quotes from the vendors ASA, Dell, Fine Tec, and SGI, together with a summary spreadsheet, are included in the Supplemental Documents. Three of them (ASA, Fine Tec, Dell) are within ten percent of the total cost of the hardware budgeted here – \$1.04 M. Their similarity in total cost and specifications gives us confidence that the estimated hardware budget, \$1.04 M, is realistic given our requirements. The fourth (SGI) is significantly higher and was discarded. The Dell and Fine Tec quotes differ in the number of GPU-accelerated nodes and assume a total storage of about 200 TB (raw).

The actual vendor and quote will only be selected after extensive product review and interviews once the proposal has been funded and the actual budget and time line are better known. Our previous experience was that vendors are more prepared to make a deal when cash is in hand. As the new machine will be assembled from components available in late 2012, it will probably differ in details from the quotes above. In particular, we anticipate the availability of faster processors with more cores, faster GPU units, and higher capacity mass storage solutions for the same cost. All these factors make our goal of a 1 PB storage facility reachable, and 0.5 PB guaranteed even in today's market. In the event that cost constraints prohibit us advancing beyond 0.5 PB, we will seek additional funding from other sources in order to advance to 1 PB.

Our storage capabilities will be built over a two year period. We have budgeted a total of \$140,000 in the second year (\$110,000 from NSF and \$30,000 from cost sharing) for storage-related hardware. In particular, we plan to add 10 TB of SSD (solid state drives) for caching and scratch space in the 2nd year of operation. This will greatly increase the performance of the overall storage system at a relatively modest cost. SDD storage will be deployed the 2nd year as SSD technology is progressing at an extremely rapid pace, and delaying purchase will inevitably buy us more SSD for the same money.

A cluster of the proposed size and complexity will require, at least initially, a full time person to look after it. A knowledgeable system manager is a vital component in helping to get the quotes, to select the vendor from the quotes, and to supervise the actual buying, installing, and setting up the HPC cluster. Furthermore, once up and running initially, queuing software will need to be developed and implemented, the operating system and RAID arrays will need to be tuned and debugged, security and licenses must be installed and maintained, initial users will need help, hardware will fail, etc, and so a system manager is essential. We have budgeted \$260 K (1.00 FTE at the level of Associate Project Scientist II for year 1 and 0.354 FTE for year 2) from matching funds to cover the (non-hardware) costs of operations over two years.

We include, in the Supplemental Documents Section, a copy of the letter from the Vice Chancellor for Research at UCSC detailing their support for this project. VCR Margon vouches that the University will provide space, power, and air conditioning as is necessary to operate the machine for its lifetime.

6 Facilities, Equipment, and Other Resources

Our proposal is to purchase and operate a new facility. The relevant existing resources are thus the infrastructure for that facility, office space for the users and administrator, and other tools for amplifying the operation and scientific yield of the new computer.

- *The site and site preparation:* The new machine will utilize the space, power, and cooling currently being used by the one it is replacing, *Pleiades.* This is an area in the basement of the Information and Technology Services Building on the UCSC campus. The conditioned and "interruptable" power is 100 kW and the cooling is 25 tons. The floor space is approximately eight racks. When Pleiades was installed UCSC spent \$450,000 upgrading this space and purchasing the UPS. The infrastructure is in perfect working order, and it has been guaranteed to us for the useful life of the new machine. It is also worth noting that the electrical power being paid for by UCSC will cost, in five years, much more than the purchase price of the cluster.

- *Ethernet connection to the outside world:* Two years ago, UCSC greatly upgraded his high speed interconnect with the outside world. This "Dark Fiber" infrastructure (shown in the Figure), provides 10 Gbps layer 3 circuit to CENIC and thus a superfast pipe to other major reseach universities and national labs. The new computer will be directly plugged into this network allowing it high speed access to our data at the national centers.

- *Visualization:* A remarkable visualization facility – the OptIPuter – is housed in CITRIS in the Scool of Engineering but is available to the whole campus. The OptIPuter is a distributed computer system capable of fulfilling many computing roles. It incorporates an optical networking computational cluster driving a tiled display wall. It is located in the UCSC Engineering Building, and boasts a display wall consisting of 40 HP monitors, each with 29.8 inches viewable area and 2560 x 1600 resolution. The displays are laid out in a 10 x 4 grid, leading to a total display size of 272 inches by 71.6 inches and resolution of 25600 x 6400 (163.8 Megapixels). This is an ideal place to display our results from multi-dimensional simulations, both for purposes of data analysis and public outreach. Many of our simulations from Pleiades have made use of this facility. See http://www.iti.ucsc.edu/optiputer.

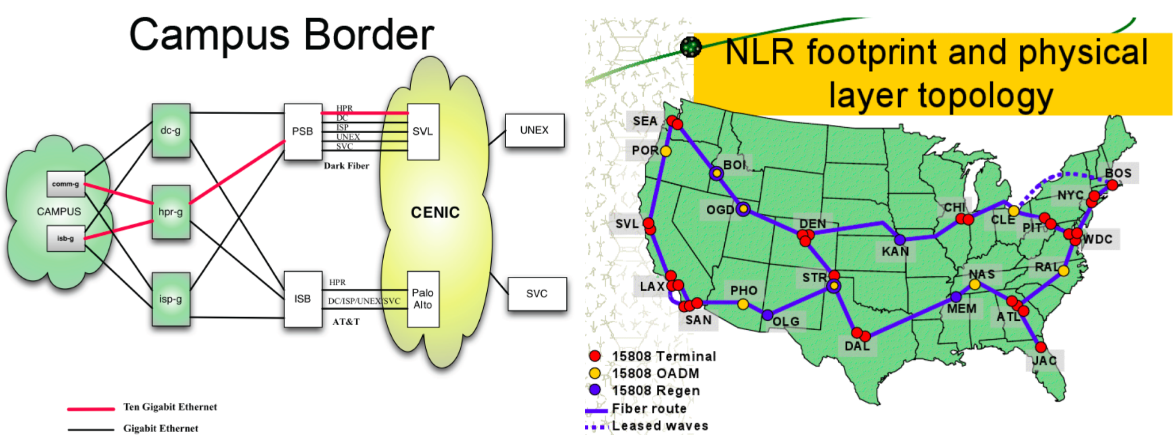

Figure 12: The UCSC Dark Fiber infrastructure, providing 10 Gbps layer 3 circuit to CENIC and thus a superfast pipe to other major reseach universities and national labs.