## 1. Astro Computation Working Group

The "Astro Computation" working group met on October 10, 2012 to discuss the intersection of the astronomy and astrophysics topics

Stellar Evolution
Core Collapse Supernovae
Type Ia Supernovae
Novae and other Transients
Nucleosynthesis
Injection of the Elements
Chemical Evolution

with

High-Performance Computing
High-Memory Computing
High-Throughput Computing
Hadoop Map-Reduce Computing
Community-Driven Social Computing
Big Data Initiatives
Next Generation Internet Capabilities

to address the questions

1) What will advanced computing and networking be able to do for nuclear astrophysics in the coming decade?
2) What do we need to do to maximize the impact of new computing and networking capabilities on nuclear astrophysics?

## 2. Summary

The dramatic impact of computation on astronomy and astrophysics is manifested in many ways. Modern numerical codes are now being used to simulate and understand the evolution, explosion, and nucleosynthesis of stars, how the elements are injected into the interstellar medium, molecular clouds, and extant planetary systems, and the cosmic evolution of the abundances. They are also essential to processing astronomical spectral databases whose sizes now exceed one terabyte into abundance data that are usable by the nuclear astrophsics community. The largest codes may have in excess of a million lines and run on supercomputers that have more than 100,000 cores, generating datasets that occupy one to a hundred terabytes of storage. The most widespread codes may be driven by comunities with hundreds of users and run on desktop class machines that generate a significant fraction of the published literature. Such codes are now an indispensable part

of the nuclear astrophysics enterprise. They often deploy teams − astronomers, astrophysicists, computer scientists, visualization professionals, applied mathematicians, and algorithm specialists − to create, maintain, and constantly develop them.

NSF, NASA, and DOE have made substantial investments in the advanced computing and networking ecosystem over the last few decades, from national to regional to university to individual facilities. Sustained peta-scale, and soon exa-scale, computing capabilities will be avalaible to the nuclear astrophysics community. Such capabilities will enable cutting-edge theoretical calculations and analyses that push the nuclear astrophysics frontier. One example is 3D core-collapse simulations run to  500 s to better quantify the neutron star winds, the r-process signatures, and evolution of the proto-neutron star. Another example is routinely deploying reactions networks with 1000's of isotopes in all 2D or 3D models. Future progress in advanced computing for nuclear astrophysics will come from further parallelization, ubiquitous deployment of next-generation 100 GB/s internet connectivity in tandem with Globus Online, distributed cloud storage systems, extracting actionable knowledge from big data, and, potentially, social computing.

Similairly, Spectroscopic stellar surveys that are scheduled over the next 5-10 years, or are recently completed (SEGUE, LAMOST, APOGEE, HERMES, GAIA, a variety of LSST followups, and the proposed 10 meter spectrographic survey telescopes), as well as all-sky surveys that will provide signifcant new information on supernovae and other transient events (LSST, SkyMapper). Nuclear astrophysics will have terabytes of spectroscopic data and petabytes of photometric all-sky data that need to be analyzed and compared to simulations.

These new technological capabilities and data driven science will enable qualitatively new physical modeling in topics relevant to nuclear astrophsyics. Exploiting these new capabilities for nuclear astrophysics will require new software instruments (e.g., run-time visualization for 100TB of 1PB data sets) and sustained funding support for focused multi-institutional research collaborations.

## 3. Reccomendations

1) Make long-term investments in appropriately focused research collaborations and codes that can make be uniquely effective in tackling some of the most difficult problems in modern nuclear astrophysics. The collaborations would be devoted to a specific nuclear astrophysics problem or topic that is believed to be ripe for a breakthrough within five years. One example, would be deploying 2D or 3D hydrodynamic simulations of the $^{13}$C pocket in low- and intermediate mass stars to produce a breakthrough in quantifying the main s-process.

2) Encourage the astro computation community to maintain common and clearly-documented data formats and standards. This will facilitate direct collaborations as well as encouraging the data to be publicly available, and will allow analysis codes to accomodate multiple data sources. For example, current challenges exist in reaction rate databases, nucleosynthetic yield data from stellar evolution calculations, and supernova simulations as input into interstellar or molecular cloud

mixing calculations or large-scale chemical evolution simulations.

3) Motivate the creation of "data libraries" for simulation inputs and outputs. This maximizes the reach and impact of the simulations, since the data can be useful for more science than the authors originally intended, thus increasing the science-per-dollar (or cpu-hour) of the simulations.

4) Inspire multiple collaborations to maintain their simulation and analsyis codes as open source projects. Having open source access to a variety of codes that do the same type of simulation or analysis is a necessary part of the scientific process. This improves transparency, cross-portability, and trust in the simulation results. For example, in a few subfields of astrophysics some results from non-open-source codes are starting to be dismissed as unreliable, and this early trend may propagate to other sub-fields of astrophysics. In addition, open source instruments lowers the barrier to entry for the next-generation of researchers and can significantly increases the amount of science produced by a community (e.g., the MESA or GR1D projects).

5) Encourage the development of an open source, community driven and supported radiation transfer code. This is a critical, but missing, piece of insfrstructure that connects stellar models with the light curves and spectra obtained from observations.

6) Persuade collaborations to engage in vigorous code comparisons and verifications as a mechanism to improve the fidelity of the codes and build trust in the results. Comparison to observations requires trustworthy simulation results. For example, the puzzeling variation in the nucleosynthetic yields from AGB stars hampers galactic chemical evolution studies. This can often be traced back to assumptions about the treatment of convection and mixing, but in some cases can be attributed to a discovered numerical instability.

7) Encourage policies that balancing capacity computing with cabability computing on the largest supercomputers. Capability computing uses all or a large fraction of the supercomouter to solve a few large problems (i.e., hero calculations). In contrast, capacity computing uses the supercomputer to solve a large number of smaller problems (e.g., surveying a parameter space). Many, but not all, members of the working group th ebalance was scewed too much towards capability computing.

8) Develop and implement an annual Summer School on numerical algorithms, parallel techniques, Big Data, and advanced computing in nuclear astrophysics.

## 4. Big Data

The phrase "Big Data" refers to large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, computational models, images and/or all other digital sources.

Big Data in nuclear astrophyiscs aims to advance the core scientific and technological means of managing, analyzing, visualizing, and extracting useful information from large, diverse, distributed and heterogeneous data sets so as to:

- accelerate the progress of scientific discovery and innovation in fields of broad interest to nuclear astrophysics;

- lead to new fields of inquiry via intellectual fusion of fields of interst to nuclear astrophysics that would not otherwise be possible;

- encourage the development of new data analytic tools and algorithms;

- facilitate scalable, accessible, and sustainable data infrastructure across the nuclear astrophysics community's efforts;

Today, the federal funding agencies and private enterprise recognize that the scientific and engineering research communities are undergoing a profound transformation with the use of large-scale, diverse, and high-resolution data sets that allow for data-intensive decision-making, at a level never before imagined. New statistical and mathematical algorithms, prediction techniques, and modeling methods, as well as transdisciplinary approaches to data collection, data analysis and new technologies for sharing data and information are enabling a paradigm shift in scientific investigations. Advances in machine learning, data mining, and visualization are enabling new ways of extracting useful information in a timely fashion from massive data sets (e.g., 3D simulations and spectroscopic surveys), which complement and extend existing methods of hypothesis testing and statistical inference. As a result, a number of federal funding agencies are developing big data strategies to align with their missions.

Embracing a Big Data initiative will help to accelerate discovery and innovation in nuclear astrophysics. The pipeline of data to knowledge to action has tremendous potential for transforming all areas of scientific interest to nuclear astrophsyics. This initiative will lay the foundations for engaging enterprise level Big Data infrastructure projects, workforce development, and progress in addressing the complex, transdisciplinary grand challenge problems in nuclear astrophysics. The field's state-of-the-art research is increasingly data-intensive and adequate sustained support for a Big Data initiative is imperative if the field is to realize its research aspirations.