# **White Paper:** Thoughts on The Future of NSF-supported Advanced Computing from a Numerical Relativity and Computational Astrophysics Perspective

Christian D. Ott (cott@tapir.caltech.edu), Mark A. Scheel (scheel@tapir.caltech.edu)
Theoretical Astrophysics Including Cosmology and Relativity (TAPIR)
California Institute of Technology

## 1 The Science and The Challenges

### 1.1 Numerical Relativity: Facilitating NSF's Advanced LIGO Project

The NSF-funded Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) will soon directly detect gravitational waves (GWs) from cataclysmic events such as the collision of two neutron stars or black holes in distant galaxies, thus testing a fundamental feature of Einstein's Theory of General Relativity. The upgraded aLIGO instruments will come online in a matter of months, and the first detection may occur this year. Numerical Relativity (NR), the numerical solution of Einstein's equations without simplifying assumptions, is the only reliable way of computing the GWs predicted by General Relativity for events like black hole collisions. So for these events, NR is the only way of generating GW predictions to be compared with aLIGO data. Thus, *NR computation is essential for making sense of data from the* $>$ *\$600M aLIGO project*.

**Looking forward: Computing needs in 2017-2020.** For black hole collisions, we estimate that 1000s of NR simulations are needed so that aLIGO can reliably test General Relativity and measure the masses and spins of the black holes. Although requirements vary greatly depending on the configuration simulated (for example, collisions of black holes with widely different masses and with large spins are much more computationally expensive), a typical simulation today requires on the order of 300k CPU-h on XSEDE/TACC Stampede. Therefore, we anticipate $\sim$ 1 billion CPU-h just for GW signals from black hole collisions.

### 1.2 Computational Astrophysics: Conquering the Multi-Physics Frontier

Massive computation is driven by the search for a fundamental understanding of astrophysical phenomena and processes that (*i*) cannot be understood from observations and phenomenology alone and (*ii*) are too rich, complex, and non-linear that analytic theory fails. Example problems include the formation and evolution of galaxies and galaxy clusters, the formation of stars and planetary systems, stellar collapse and explosions (supernovae), and the kinetics and dynamics of astrophysical plasmas and shocks. These are all multi-scale, multi-dimensional, and multi-physics problems. Only computation can provide a fundamental understanding of these processes/phenomena and can make theoretical predictions that can be tested with observations.

In all areas of computational astrophysics the challenge is now to find computationally tractable ways of including all the (known) necessary physics for the solution of a given astrophysical problem. This requires access to massive advanced computing resources, but also necessitates novel algorithms and approaches to fully exploit current and future resources.

**Looking forward: Computing needs in 2017-2020.** Computational needs vary with the problem at hand but are steadily growing as more physics is incorporated in the simulations. As an example: a single end-to-end 3D core-collapse supernova simulation through the first second of the explosion requires $\gtrsim$5 million node hours on Blue Waters ($\gtrsim$50 million service units on TACC Stampede). Many of these simulations, with higher resolution than currently practical, are needed to ascertain the still uncertain explosion mechanism of massive stars and to study the final outcomes of these explosions and their nucleosynthetic imprint on the Universe.

Other problems, e.g., star formation and galaxy evolution, have similar or even greater computational needs that must be met to leverage upcoming observational efforts such as the Large Synoptic Sky Survey (LSST) and 30-meter-class telescopes. Massive computations will be needed to interpret and understand the data obtained with these instruments.

### 1.3 Algorithmic Challenge: Disruptive Changes are ahead.

The transition from current petascale to future exascale architectures will be a disruptive change. *Current capability computing problems will not simply be capacity computing problems in the exascale era*. Unless rewritten, (relativity/astrophysics) codes will simply not scale and execution will be slower than on existing petascale systems. Most codes are unable to achieve more than a few percent of current theoretical peak performance and are not ready to benefit from available accelerator technology. The more physics that is included, the greater is the challenge to scale even to petascale. It will be a tremendous technical and algorithmic challenge to re-tool these codes for exascale. New parallelization and task scheduling paradigms and resilient, fault-tolerant algorithms will be essential in the exascale era – even for codes trying to solve petascale problems.

## 2 Discussion Points

### 2.1 Big Data is Not Everything. It would be bad for NSF to focus on Big Data alone.

The rapidly improving ability to collect, classify, store, and analyze data is having a massive impact across all scientific disciplines. But as impressive and revolutionary as Big Data is, it is ultimately an extension in scale of *one key piece* of the scientific method: observation and experiment.

The other, equally important piece is theoretical modeling and computation. *Big Data-driven science is complimentary to but cannot and will not replace theory and computation. And in many cases, Big Data requires massive computation to be useful*. For example, GW data from aLIGO and particle physics data from the LHC are most useful in conjuction with theory-based computational simulations of potential experimental outcomes that can be tested by the experiment. Data from the upcoming Large Synoptic Sky Survey and the Square Kilometer Array will likely revolutionize our observational picture of the Universe, but they will not *on their own* provide fundamental explanations for the nature of dark energy and dark matter, or the dynamics of galaxy evolution.

The Sloan Digital Sky Survey (SDSS) was the first true Big Data project in astronomy and amassed $\mathcal{O}(100)$ TB of data. It has indeed driven our understanding of the Universe, but it has in no way replaced the need for theory and cosmological simulations. In fact, SDSS has increased the need for and the complexity of such cosmological simulations.

Big Data will not render theory and computation obsolete – *Big Data will require more theory and computation.* It would be disastrous for both fundamental and applied science if NSF decided to focus its support of advanced computing only on Big Data.

### 2.2 Big Compute does not mean No Big Data.

Big Data and Big Compute can both cost-effectively make use of common computational infrastructure. Current advanced computing systems such as TACC Stampede and NCSA Blue Waters are designed so that node-level and rack-level heterogeneity can be supported. For example, Stampede has high-memory nodes for memory-intensive and data-intensive tasks. Other specializations within a common global platform are possible: e.g., fast node/rack-local solid state disks, ultra-fast storage interconnect, specialized accelerators (GPUs, FPGAs) etc. can all be integrated into current and future HPC solutions and can have cross-functionality between data-driven and compute-driven applications. *Big Data and Big Compute are not two diametrically opposed paradigms. They can and should coexist on the same platform.*

### 2.3 NSF should mandate Open Source Scientific Software.

In many areas of computational physics and astrophysics, proprietary scientific software is the norm and open source software is the exception. Because the scientific publication process focuses on results rather than methods, many proprietary scientific codes are undocumented or incompletely documented in the literature. **This makes many computational studies effectively irreproducible,** violating one of the fundamental aspects of science.

Examples show that open source software accelerates scientific progress while proprietary software leads to its stagnation. The availabilty of open source cosmological hydrodynamics codes like *Gadget* has lead to many breakthroughs in our understanding of how the Universe and galaxies evolve. The absence of open-source codes for studying core-collapse supernovae may be one reason why the study of these objects is dominated by very few research groups who have still failed to explain how massive stars explode after five decades of effort.

While NSF-funded MREFCs must release data to the public, and individual proposals must include a data management plan, there is no equivalent for software. The fact that there is no mandate for software to be released as open source violates the paradigm that publicly funded research should lead to publicly accessible results. Requiring open source would increase science return per dollar spent, because: (A) Different researchers would not need to develop equivalent proprietary technology from scratch. (B) Code would be more reliable (more eyes on the same code) and results reproducible. (C) Researchers with disruptive new ideas could more easily enter established fields and drive scientific progress.

### 2.4 Campus Computing is a crucial component of the national Cyberinfrastructure.

Campus and department level computing on clusters with 1–10k cores plays a crucial role in the US advanced computing ecosystem. These Tier-2 systems are crucial for code development, rapid-turnaround testing/debugging, smaller-scale capacity jobs, and staging for Tier-1 capability machines. *It is impossible to go from a laptop/desktop to a Tier-1 computer without access to Tier-2 resources*.

NSF's current approach to funding of campus-level Tier-2 resources is through the very broad and general Major Research Infrastructure (MRI) program. Given the increasing demand for computing across all disciplines (including the social sciences), *a dedicated ACI program specifically for support of Tier-2 campus-level computing is needed.* This should replace the current MRI program that puts interdisciplinary campus/department level computing in competition with very specialized scientific hardware such as electron microscopes and spectrometers.

### 2.5 US Science needs NSF-supported Access to Advanced Computing.

No other federal agency provides the same broad access to advanced computing resources as NSF. NSF's lack of mission focus (unlike DOE and NASA) is its true strength in providing democratic merit-based access to computing for all disciplines. *The current NSF XSEDE model should be preserved under all circumstances*. Synergies and reduction of cost by partnering with DOE or NASA should be exploited wherever possible, but the NSF's share of future joint cross-agency machines should be allocated in the proven effective way. The broad access to advanced computing resources through XSEDE has been crucial for US leadership in computational science. *NSF should not try to fix a process that is not broken.* Congress, through NSF, should provide more and updated large-scale capability and capacity/throughput compute resources. XSEDE resources are currently dramatically oversubscribed and major investments are necessary. Any trade-off on advanced computing will hurt US competitiveness across all scientific disciplines.

### 2.6 Instrument (= Code) Scientists are crucial and we must support them.

Scientific (open source) code should be regarded as a *software instrument* and discipline scientists who develop and maintain code should be regarded, credited, and valued as *instrument scientists*. DOE would not build a particle accelerator and then eliminate the position of the instrument scientist who knows how to operate and fix it. In the same sense, NSF should (work with universities to) create long-term computational scientist positions. This is the way to offer the long-term perspective necessary to retain the best and brightest code-oriented computational scientists.

### 2.7 A permanent national Cyberinfrastructure to reduce Costs.

In the current paradigm, NSF supports individual large-scale computing systems that are computed among private/state-funded computing centers. There is no (federal) guarantee that these centers will exist over the lifetime of multi-decade NSF's MREFC projects like LIGO, LSST, Alma, etc. Hence, $\mathcal{O}$(10-100) million dollars are spent on MREFC-specific cyberinfrastructure that could be saved if there was a long-term stable national cyberinfrastructure in the form of permanent computing/data/software centers. Congress, through NSF, should make a long-term investment in such centers and regularly re-compete their management rather than their resources.