Fluid Dynamics and Turbulence: The Case for NSF-HPC in 2017-2020

P.K. Yeung (Georgia Tech)

Fluids are necessary for life, and an understanding of fluid dynamics is vital for major advances in aircraft design, weather and climate prediction, environmental and energy sustainability, biomedical technologies, and many other applications important to society. Although the underlying physical conservation laws are well known, the equations of fluid dynamics are mathematically intractable, especially if the motion is turbulent. The need for computation from a theoretical perspective is thus acute. Advances in computing power in the last 30 years have led to much success in numerical simulation, as the best and most detailed source of information for understanding and modeling turbulent flows. The reliance on computation is even more striking in problems such as Earth mantle convection and solar weather where experiments are impossible. At the same time, a diverse computational ecosystem has emerged, whereby many science discoveries are also made through qualitative and quantitative analyses of massive (''Big'') data even at moderate resolution. Yet, the pace of progress for fundamental research and US leadership compared to our world peers in this field are now clearly tied to the continuing availability and future growth of advanced computing resources that NSF has provided since the mid 1980s.

We focus our discussion below on turbulence, which is one of the greatest scientific challenges of all time, with diverse applications. We also consider the views and needs of the fluid dynamics community at large, in the context of their experiences as users of supercomputing facilities.

Subject-specific research needs

Turbulence occurs over a wide range of scales. The two most resource-intensive approaches are direct numerical simulation (DNS) which resolves all scales, and large eddy simulation (LES) which resolves only the large scales but models the small scales. Although numerical schemes vary, DNS at high flow speeds tends to be limited by interprocessor communication made necessary by distributed-memory parallelism. The largest DNS performed on Blue Waters at UIUC uses over half a trillion grid points and over 200,000 CPU cores. Many researchers are also working on LES applied to complex geometries, and problems where other processes such as particles and droplets, density stratification and chemical reaction also arise. A realistic simulation of rain formation, involving atmospheric thermodynamics, phase change, particle inertia, humidity, and radiative heat transfer, would be a worthy Exascale problem that has very significant impact. It is clear that fluid dynamics has a leading role in this endeavor.

Impact of academic NSF advanced computing resources on subject research

NSF grantees funded by the CBET Fluid Dynamics program are known to be more active in requesting and using XSEDE resources than their peers associated with the Engineering Directorate. The tremendous scientific impact of NSF investments in advanced computing for fluid dynamics is undeniable and readily seen in the acknowledgment sections of numerous papers in leading journals of our field. For many of these investigators XSEDE facilities are their natural and only choice for computations that are too large to be carried out using clusters available on their own campus (which would, in contrast to the economy of scale provided by the large centers, also be much more costly in aggregate). This is particularly important for researchers working on fundamental and scholarly problems, where demand for resources often outweighs available funding.

Role of other agencies and national laboratories

Scientists pursuing mission-driven research are generally able to obtain time at centers of various sizes directly supported by the funding agencies. The INCITE program at the DOE's Argonne and Oak Ridge National Laboratories provides access to Leadership Class facilities but only gives a limited number of awards and is dominated by large multi-institution teams sometimes involving DOE laboratory scientists. As such, this program is not suitable for researchers focusing on medium-sized problems, and those working in areas with no or little direct connection to the DOE's mission have found it harder to compete. While the DOE has already announced plans for new machines past 100 Pflop/s, there is a risk that fundamental fluid dynamics could be left behind.

Resource allocations: supply, demand, and review processes

It is well known that demand has generally outstripped supply. This situation may be particularly acute at this time since only one XSEDE machine is currently at or above the 1 Pflop/s mark. XSEDE resource managers (and their TeraGrid predecessors) have handled the process conscientiously, but have had to make across-the-board cuts of almost 50% even on proposals receiving high ratings. The review process is currently quite transparent and there are sufficient safeguards against conflicts of interest. However, sometimes because of mismatch of interests some reviewers may not fully appreciate the scientific impact of the proposed computations. Multi-year resource allocations would give many researchers more flexibility but such requests appear to have a low success rate perhaps because long-term planning of computational details is very uncertain.

Given these challenges, a growing number of investigators have sought allocations from sources abroad (mainly Europe, through international partners, joint appointments, etc.). Such a "silicon drain" out of US science, especially at the mid-levels, is in fact quite alarming. It is important that NSF help us and other science communities avoid slipping behind on the international scene.

Users' point of view on facilities at the national centers

XSEDE and formerly TeraGrid centers have provided good service, in key aspects such as training, documentation, consulting, day-to-day operations, job scheduling, and mass storage. Unscheduled interruptions and system issues are unavoidable but have been handled professionally. However sometimes the introduction of new machines of markedly different architectural properties have caused concern in the learning curve required to reap the full benefits of the new functionality. While large research groups (not necessarily the most resource-thirsty ones) may be able to devote significant manpower where necessary, others may struggle to strike a balance between time spent in writing and optimizing codes versus the time needed for the pursuit of science itself. A second concern is long-term data archival, since data continue to have value for years after resource allocations and grants expire, in fact even when or if a HPC center itself loses funding. The volume of data concerned is often far larger than what an individual researcher's institution can accommodate or process. This is an area where some deep policy discussions are required.

Recommendations for NSF-HPC, 2017-2020

Since the mid-1980s NSF-HPC has played a phenomenal role in US science leadership. Any weakening of this support will have disastrous effects on our and other science communities. Despite hardware and budgetary challenges, this support must continue. We recommend that NSF:

1. Invest in a future hierarchial Cyberinfrastructure portfolio that will both enable Exascale computational science by 2020, and provide adequate benefit to the computational middle class and data-intensive applications. Closer links between HPC and Big Data are also needed.
2. Develop interdisciplinary grant programs that will help and encourage researchers to benefit fully from new systems and to pursue scientific discovery at the highest levels. We also advocate a closer advisory role for NSF disciplinary program directors in HPC governance.

This position paper has been prepared with help from several colleagues, namely: L.R. Collins (Cornell), S. Elghobashi (UC Irvine), C. Meneveau (Johns Hopkins Univ), M.W. Plesniak (George Washington Univ), W.W. Schultz (Univ. Michigan), and K.R. Sreenivasan (New York Univ).