

NERSC's Impact on Advances of Global Gyrokinetic PIC Codes for Fusion Energy Research

Stephane Ethier, Choon-Seock Chang, Seung-Hoe Ku, Wei-li Lee, and Weixing Wang |

Princeton Plasma Physics Laboratory

Zhihong Lin | University of California, Irvine

William Tang | Princeton University

The National Energy Research Scientific Computing Center (NERSC) was originally launched as a computing center for the exclusive support of magnetic confinement fusion research in the US. One example of the numerous computational achievements enabled by NERSC is the development of the global gyrokinetic particle-in-cell approach for the simulation of turbulent transport in tokamak fusion devices.

The fusion energy research community has always had a special relationship with the National Energy Research Scientific Computing Center (NERSC). Historically, NERSC started as a dedicated computing center for fusion research, first as the Controlled Thermonuclear Research Computing Center at the time of its inception in 1974, followed by the National Magnetic Fusion Energy Computer Center in 1976. Since 1990, NERSC serves all of the science domains funded by the US Department of Energy's (DoE's) Office of Science, although its impact on fusion energy research has undeniably been among its greatest. The vast amount of computer time delivered by NERSC over the past four decades has been crucial to the development and advances of magnetic confinement fusion science, including the science of global gyrokinetic particle-in-cell (PIC) codes.

The idea of confining a 100-million-degree ionized gas, called plasma, using only magnetic fields might seem surprising at first, but it's presently the most successful and efficient way of producing fusion energy in a controlled way. Of all the "magnetic bottle" configurations built and tested since the start of fusion research in the 1950s, the

doughnut-shaped "tokamak" (see Figure 1) continues to be the most successful design. The torus is the only shape that ensures a nonvanishing magnetic field everywhere in the volume while maintaining a very large fraction of closed field lines. These are essential requirements for confining the charged particles forming the plasma as they follow the magnetic field lines while executing cyclotron motion.

With the vast amount of thermodynamic "free energy" available in a 100-million-degree plasma with steep pressure gradients, extending from the hot dense core to the "cold" low-density wall, turbulence and instabilities can quickly develop under certain conditions. The time and length scales of the important phenomena in fusion plasmas span 10 orders of magnitude, from the fast (approximately 10 GHz) and short (roughly 10 μm) electron scale to the slow (few seconds) and large (few meters) scales of the entire system. Because of this, several computer codes are needed to treat the physics at the different scales. A full-wave electromagnetic solver with kinetic description of the plasma is used to calculate wave heating of the plasma on the fast electron and ion cyclotron frequencies. Low-frequency turbulence, which is believed to be the

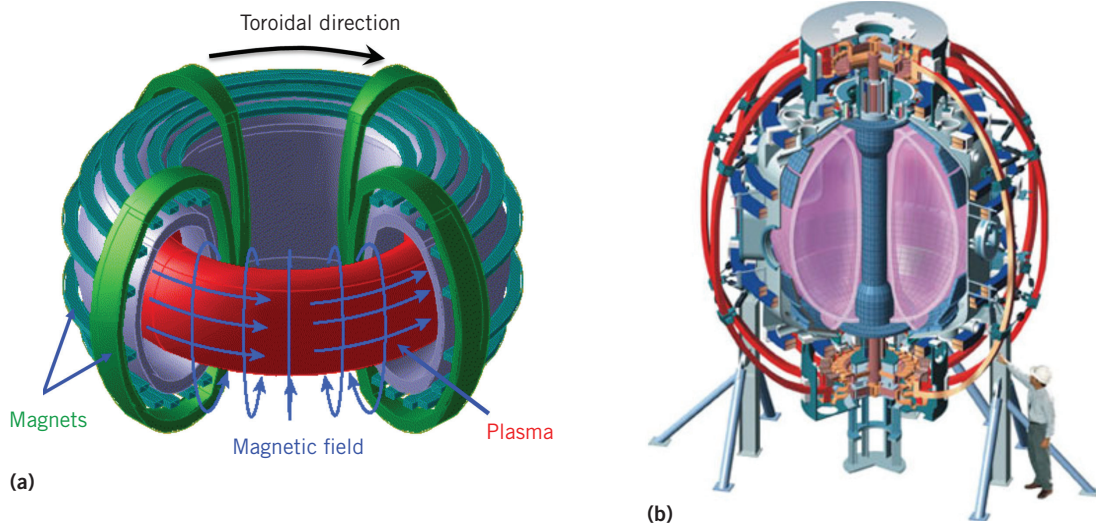


Figure 1. Schematic of a tokamak magnetic confinement device: (a) a generic tokamak and (b) the National Spherical Torus Experiment (NSTX) at Princeton Plasma Physics Laboratory. The latter is a strongly shaped tokamak with a small aspect ratio.

main mechanism by which the plasma loses energy via cross-field transport, also requires a kinetic description of the fusion plasma, but at a time scale much longer than the ion cyclotron period, although not long enough for the core plasma profiles to change significantly. The large system-size instabilities are normally simulated using a magneto-hydrodynamic (MHD) code, which evolves a set of fluid equations appropriate for a plasma subjected to a strong guiding magnetic field. Plasma profiles and magnetic fields evolve significantly on this time scale. Finally, a so-called “transport” code is used to simulate the whole tokamak discharge on the full device time and length scales. In these codes, the physics at much faster time scales is normally treated with reduced or simplified models validated through experiments and from the more detailed codes just mentioned. A great number of computational applications have been developed and run at NERSC over the years to support the magnetic fusion energy (MFE) program. DoE laboratories, such as the Princeton Plasma Physics Lab (PPPL), Lawrence Livermore National Lab (LLNL), Oak Ridge National Lab (ORNL), and Los Alamos National Lab (LANL), have all made important contributions to the development of MFE, as have many universities and a few companies (such as General Atomics) across the US. Of all these institutions, PPPL is unique because it’s almost entirely dedicated to the advancement of fusion as a viable source of clean energy.

Most of the MFE-related codes developed at PPPL have relied on NERSC for compute cycles and computing know-how (programming, optimization, and so on). State-of-the-art MHD codes have been developed and continue to be developed at PPPL to address issues such as plasma stability, prediction of disruptions in tokamaks, and the impact of “sawtooth” instabilities. One type of application, however, has deep roots at PPPL and has always accounted for a large amount of computer time at NERSC: the global gyrokinetic PIC code. The remainder of this article describes the evolution of this type of code at PPPL and how it has been closely tied to NERSC.

Beginnings of Plasma PIC

One of the pioneers of the PIC method and a father of plasma computer simulation was a plasma physicist at PPPL from 1956 to 1973.¹ John Dawson left a strong legacy of investigating fusion plasmas with particle codes at PPPL before moving to the University of California, Los Angeles (UCLA), to form a new simulation group that went on to develop new ideas not only in fusion plasmas but also in accelerator physics using state-of-the-art PIC codes to study plasma-based acceleration, an original idea from Dawson about “letting particles surf on the plasma-wave wakes left behind by a laser or a particle beam as it moved through plasma” (<https://plasmasim.physics.ucla.edu/research/dawson>). However, Dawson wasn’t the only one who helped

develop the PIC approach. The fundamentals of the method were laid down in the late 1960s and early 1970s by a few small research groups located at the University of California, Berkeley,² Princeton,³ UCLA,⁴ LLNL,⁵ LANL,⁶ Stanford University,⁷ the Naval Research Lab,^{8,9} and the UK.¹⁰ The two most cited textbooks describing the PIC approach are by Charles K. Birdsall and Bruce Langdon,¹¹ and Roger W. Hockney and James W. Eastwood.¹² Although Dawson left PPPL, the development and use of PIC codes for studying MFE continued to grow due to the work of researchers such as Hideo Okuda and Wei-li Lee. Both started using NERSC soon after the start of its operation in 1974; in fact, most PIC pioneers in the US became NERSC users.

The PIC algorithm takes advantage of Liouville's theorem, which states that phase space, the extended 6D space of positions and velocities, is incompressible for a Hamiltonian system. Particles are used in the simulation to sample the phase space in a Monte Carlo sense, each one being evolved along the characteristics using a Lagrangian approach while ensuring the incompressibility of phase space. Because of this, a small volume of phase space moves along with each particle during the whole simulation, carrying with it the physical properties of that volume (6D space size and density) wherever the particle goes. These properties are used to calculate physical, macroscopic quantities to be compared with experiments. In plasmas, the particles are ions and electrons, which are subjected to long-range Coulomb interactions due to their charge. For the interaction of each pair of charged particles, this leads to an n^2 calculation, where n is the number of particles, including both the long-range (small deflections) and short-range (large deflections) interactions. The long-range interactions describe the collective effects in the plasma, which lead to waves and instabilities, while the short-range interactions result in angular deflection collisions. To avoid the very costly n^2 calculation, the PIC method uses a grid to treat the interactions between ions and electrons. The charge of each particle is accumulated on the nearest grid points around that particle. Several interpolation functions, called *shape functions*, are possible, although the linear interpolation is the most common.

Once all the charges have been summed up on the grid, the electromagnetic equations (Maxwell's equations) are solved using a grid-based solver, which is usually a discrete Fourier transform

(DFT) solver. The calculated forces are then interpolated back to the particles, which can now be advanced for another discrete time step using an ordinary differential equation (ODE) solver, such as the Leapfrog or Runge-Kutta methods. The procedure is then repeated with the new positions of the particles. The use of a grid not only gives a computational advantage by reducing the number of operations from n^2 to n , but it also delivers a fully self-consistent system that treats the long-range collective interactions between the charged particles of the plasma without overwhelming the simulation with the large fluctuations that can arise from close-range interactions. All interactions at a range smaller than the grid spacing are "smoothed" out by the charge accumulation interpolation. Because of this, the grid spacing must be chosen consistently with the physics being simulated. The classic PIC algorithm uses the Debye length as the grid spacing, which is the distance beyond which the electric field of a charge is screened by the other charged particles. That distance is fairly small for a typical plasma (approximately $10\ \mu\text{m}$), requiring a large number of grid points to correctly simulate a laboratory experiment. However, many important phenomena can be demonstrated and explained with a small system using a modest number of particles and grid points. For example, the PIC algorithm is ideal for simulating kinetic effects such as Landau damping and the two-stream instability.

Development of Gyrokinetic PIC and Delta-f Methods

Although "classic" PIC is a very powerful method for handling nonlinear physical systems that require kinetic treatment, the time step used in the simulations is determined by the highest frequency waves in the system. In a tokamak, these are the plasma waves (approximately 10^{-9} sec for ions and 10^{-11} sec for electrons) and cyclotron waves (approximately 10^{-7} sec for ions and 10^{-10} sec for electrons). A time step too large compared to the characteristic period of these waves leads to an unstable algorithm when an explicit time-advancing scheme is used. Because of this, applying the conventional 6D+time PIC method to a full tokamak discharge is still out of reach today, although exascale computing could change that in the coming years.

In the late 1970s/early 1980s, one of us (W.W. Lee) developed the gyrokinetic approach for particle simulations for the study of low-frequency microinstabilities in magnetically confined plasmas.¹³

These instabilities are responsible for the energy and momentum losses across the confining magnetic field lines in a tokamak. Understanding their effects and developing ways to minimize their impact have been the focus of intense research in the past few decades. The effective frequencies of these waves, called *drift waves*, are much lower than the characteristic frequencies of the plasma and cyclotron waves. In a tokamak's strong guiding magnetic field, the ions and electrons describe a fast spiral motion around the magnetic field lines due to the Lorentz force. This cyclotron motion is much faster than the instabilities' growth rates, so resolving that motion isn't necessary and is actually prohibitively expensive. However, the size of the ion cyclotron orbit is comparable to the wavelength of the instabilities perpendicular to the field lines, so it's important to resolve that scale length. The solution to this problem comes in the form of the gyrokinetic equation, which is a mathematically rigorous coordinate transformation of the 6D kinetic equation to the "gyrocenter" coordinates, which is the center of the spiral orbit described by the charged particles, followed by a time average (so-called gyrophase average) of the fast cyclotron motion, which reduces the kinetic equation to 5D and preserves all the nonlocal effects of the cyclotron orbits of the ions while eliminating its fast time scale. This new equation opened up a whole new field of research activities in MFE and space plasmas, which continues today.

Although some researchers were looking to numerically solve the gyrokinetic equation with grid-based methods, Lee showed that the PIC approach could be applied rigorously to solve the same equation, essentially replacing the point particles with "rings" representing each individual particle's averaged cyclotron motion. The charge accumulation step of the algorithm consists of picking four equally spaced points on the ring to obtain a valid nonlocal contribution of each charged particle to the field.¹⁴ This new algorithm, designated as the four-point average method, relaxed the time-step restriction on the particle advance, making possible a whole new field of numerical studies. By allowing for a much larger time step, fully kinetic simulations of low-frequency microinstabilities on the order of 10^{-3} seconds in total physical time were finally possible.

A few years later, Lee and a postdoctoral fellow (Scott Parker) developed a fully nonlinear delta-f version of the gyrokinetic equation¹⁵ that dramatically reduced the number of particles required for

gyrokinetic PIC simulations. This approach uses a multiscale expansion to separate the distribution function of the particles in a fixed, analytical part and a perturbed (δf) part, allowing for the first fully 3D global gyrokinetic PIC simulation of a laboratory-size tokamak in toroidal geometry.¹⁶

The Gyrokinetic Toroidal Code (GTC)

In the 1990s, parallel computers became more prevalent and so were several programming models to make use of them. While most of the PIC work at PPPL had been carried out on NERSC's vector processor Cray systems, parallel computers opened up a whole new realm of possibilities. With its 512 processors linked together by a fast network, the Cray T3E "mcurie" supercomputer installed at NERSC in 1997 was a true game changer. Zhihong Lin, a young PPPL researcher at the time, took full advantage of this new technology and developed the Gyrokinetic Toroidal Code (GTC), which introduced a highly optimized algorithm for simulating microturbulence in tokamak fusion devices. A few other groups in the US also took advantage of the new computing capability of the early 1990s to develop full-torus, global gyrokinetic codes.¹⁷⁻²⁰

In GTC, however, the combination of three key innovations stood out. The first was the use of the Hamiltonian guiding center orbit algorithm for advancing charged particles in a strong magnetic field. Developed by Roscoe White, Morell Chance, and Allen Boozer for the Orbit code,^{21,22} this algorithm casts the energy-conserving Hamiltonian equations of motion in magnetic coordinates, resulting in highly accurate trajectories because the charged particles traveling along a magnetic field line see a straight line even though that field line might be twisted in Cartesian space. The second innovation was the use of a field-line-following mesh that reduces the number of grid points by about two orders of magnitude by taking advantage of the quasi-2D structure of the drift waves being simulated (Figure 2). This is a consequence of the fast motion of the charged particles along the magnetic field lines and their much slower drift motion across those same field lines. Having a grid that follows the natural configuration of the magnetic field as it twists around the torus, combined with the use of a coordinate system described by the magnetic field itself, results in a highly accurate and efficient calculation compared to the previous method. However, the "twisted" grid removed the periodicity of the toroidal direction, which is

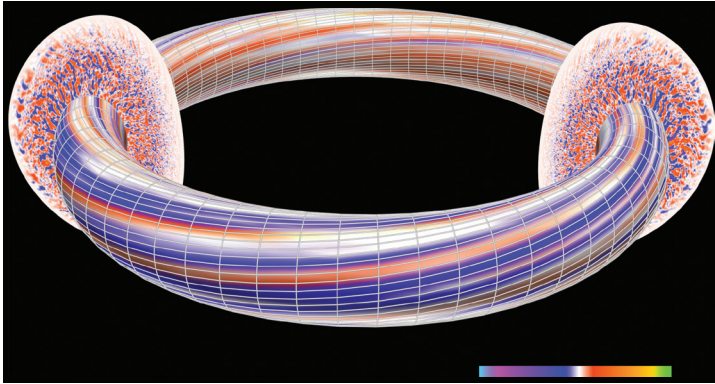


Figure 2. Snapshot of a GTS ion-temperature-gradient instability simulation showing the field-line-following mesh and the quasi-2D structure of the electrostatic potential in the presence of microturbulence. Notice the fine structure on the two poloidal planes perpendicular to the toroidal direction, while the potential along the field lines changes very little as you go around the torus. (Image generated from a GTS simulation by Kwan-Liu Ma and his group at the University of California, Davis.)

required by the spectral solver based on fast Fourier transforms (FFTs) traditionally used in PIC codes to calculate the self-consistent field. This was addressed by the third innovation in GTC, which was the development of a nonspectral solver for the gyrokinetic Poisson equation.²³ This new solver didn't require the use of FFTs and allowed for a low-communication parallel algorithm using domain decomposition in the toroidal direction instead of a full replication of the 3D grid on all processors or an expensive global transpose of the grid when using FFTs along with domain decomposition.

NERSC was also key in introducing its users to new parallel programming models, such as the message-passing interface (MPI) and "shmem," both highly efficient on the Cray T3E (although shmem had lower latency than MPI). By using MPI, the low-noise delta-f method developed by Parker and Lee, and a Boltzmann model for the electrons (so-called adiabatic electron model), Zhihong Lin and his colleagues carried out first-of-their-kind global simulations of ion temperature gradient-driven microturbulence in a real-size tokamak and showed the reduction of radial energy transport by the nonlinear development of zonal flows in the system.²⁴ These simulations used 100 million particles and 25 million grid points, and were dubbed "massively parallel" at the time, with 64 processors.

As is the case for most fusion codes, GTC was written in Fortran. However, it used the Fortran

90/95 constructs, which were fairly new at the time. Most of the parallelism and computational load come from the particles in PIC codes, even when using the delta-f algorithm, which reduces the number of particles required for the simulations. Depending on the physics being studied, the number of particles exceeds the number of grid points by 10 to 100. About 90 percent of the runtime is spent in calculations involving the particles, the two most time-consuming being the particle "push" and "charge accumulation" steps, which account for 85 percent of the time in more-or-less equal parts. The rest of the time is accounted for by the field solve, MPI communications, diagnostics, and so on.

Global Gyrokinetic Simulations

In 2001, NERSC acquired a large IBM SP computer featuring the latest Power 3 processor. While the Cray T3E was a strictly distributed parallel computer, the IBM SP was a hybrid system, comprised of symmetric multiprocessor (SMP) nodes linked by a fast interconnect. Each node contained 16 IBM Power 3 processors running at 375 MHz (1.5 Gflops peak) and sharing 16 Gbytes of memory in a symmetric fashion. Starting in 1999, NERSC provided a test system (gseaborg) to users to start porting their codes. Intensive training sessions by both the NERSC staff and IBM's Advanced Computing Technology Center (ACTC) group greatly helped users get the most out of the new supercomputer. With its one-dimensional domain decomposition in the toroidal direction, GTC was limited to about 64 processors because the field-line-following mesh made unnecessary the use of a larger number of toroidal grid points (planes). All the codes running on Seaborg at the time were pure MPI applications, using 16 processes per node and splitting the 16 Gbytes of memory (32 Gbytes after the 2002 upgrade) equally. At 1 Gbyte per process, this was already four times more than on the Cray T3E. However, all 16 processes had to share the same communication port on the node to access the network for MPI communication, which was a challenge for communication-intensive applications.

One solution was to use a single MPI task per node and multithreaded parallelism for the 16 processors sharing the memory on these nodes. NERSC and IBM recommended using the relatively new OpenMP programming model based on easy-to-use compiler directives (www.openmp.org). Stéphane Ethier had already been exploring OpenMP

on an SGI Origin 2000 at Princeton University for a different project, so implementing it in GTC was relatively straightforward. However, obtaining good performance required the specialized knowledge of the NERSC consultants and IBM ACTC group. A fine-grained multithreaded parallelization scheme was adopted because it didn't interfere with the MPI. The scatter operation in the charge accumulation step of the code was particularly challenging due to memory collisions between the OpenMP threads. The solution was fairly simple, though—give each thread its own local private grid to accumulate the charge and perform a reduction, or merge, of all the thread contributions at the end. This improvement let GTC scale to 1,024 processors on Seaborg while giving each MPI process the full 16 Gbytes of memory on each node. With this modification, we were able to run the first global, kinetic simulation of the ITER tokamak, which is the largest and most important fusion experiment currently under construction in Europe as part of an international collaboration of seven countries, including the US (www.iter.org). The simulation used 1 billion particles and 125 million grid points, which was unprecedented; it was part of a numerical study investigating the scaling of energy losses due to turbulent transport as tokamak size increases.²⁵ Our simulations showed that the energy losses level off as the tokamak size approaches ITER, a reassuring result for ITER designers.

In 2002, Seaborg was upgraded to more than double its size (6,656 cores) and twice the memory per node (32 Gbytes/node). Through a continuous increase in user code concurrency, new challenges were encountered and addressed in the following years. This knowledge, as well as the requirements of the upgraded user applications, paved the way for the new Leadership Class Computing Centers at ORNL and ANL. Many users were readily able to run on the large Leadership Class Computers due to their work at NERSC. This was certainly true for GTC, which Ethier had upgraded with a new level of MPI parallelism,²⁶ allowing for a much larger number of particles in the simulations. The particles within each toroidal section could now be distributed between several MPI processes. A reduction operation of all the processes involved took care of summing up all the contributions to the local charge accumulation. This capability allowed GTC to perform, among other things, very large simulations of electron temperature gradient (ETG) instabilities. Also in 2002, GTC's main author, Zhihong Lin, became a professor at the

University of California, Irvine, where he formed a new gyrokinetic PIC group.

GTC continues to acquire new features, such as energetic particles,²⁷ kinetic electrons,²⁸ and electromagnetic capability.^{29,30} At PPPL, in collaboration with LBNL computer scientists from the Future Technologies Group,³¹ GTC continues to be a testbed for the development of new high-performance algorithms. The MPI particle distribution scheme has been replaced by a full-fledged domain decomposition to further divide the torus and reduce the memory footprint of each MPI process. This highly scalable version, called GTC-Princeton (GTC-P), has been run on 1.5 million cores on the IBM Blue Gene/Q system at LLNL.

Gyrokinetic Tokamak Simulation (GTS) Code

Although GTC had been highly successful as a scalable, theoretical tool to uncover new phenomena related to microturbulence in toroidal fusion devices, it could only handle a relatively simple geometry consisting of a circular cross-section tokamak with a large aspect ratio analytical equilibrium magnetic field (the aspect ratio is R/a , where R is the major radius of the torus and a the minor). The temperature and density gradient profiles were also given as analytical formulas controlled by a few parameters. However, there was a strong interest in having GTC run simulations with geometry and plasma conditions as close as possible to the real tokamak experiments, especially for PPPL's own experimental device, the National Spherical Torus Experiment (NSTX), which is strongly shaped and has a low aspect ratio (Figure 1b). In 2004, PPPL researcher Weixing Wang engaged in the development of a new global gyrokinetic PIC code that could handle any tokamak shape and profile in a general and robust way (Figure 2). The result was GTS, or gyrokinetic tokamak simulation code, which interfaces with the experimental database to retrieve temperature, density, and rotation profiles along with magnetic field information for tokamak shots.³²

The higher-level structure of the code was similar to GTC, but most functions and routines had to be rewritten to support a whole set of new features—for example, different magnetic coordinates had to be implemented to keep the poloidal planes “flat” in Cartesian space, which in turn forced the use of new Lagrangian guiding center equations for the particle trajectories. The field solver became more complex, requiring the inclusion of the zonal flow in an iterative process with the local solve. A

new delta-f scheme based on extended phase space was later implemented to enforce strict phase-space incompressibility, which made the code more reliable, accurate, and robust.

All the performance and parallel scalability optimizations learned through the development of GTC at NERSC were also implemented in GTS. For example, the efficient field-line-following mesh is used, as well as all three levels of parallelism, which are the toroidal domain decomposition, particle distribution within the domains, and loop-level shared memory multithreading with OpenMP. Most of the GTS simulations have been (and continue to be) run on the systems at NERSC. Many successful and diverse studies in transport and confinement have been carried out, including impurity transport, trapped-electron-mode turbulence,³³ current generation, momentum transport and spontaneous rotation, global ETG turbulence in NSTX,³⁴ flow generation,³⁵ shear-flow-driven instability, turbulence spreading and nonlocal transport, energy transport of low-k turbulence in various experimental tokamaks,³⁶ and turbulence-driven energetic particle transport.

In 2009, GTS benefited from dedicated support funded through the LBNL Petascale Initiative in Computational Science and Engineering (<https://www.nersc.gov/news-publications/news/nersc-center-news/2014/petascale-post-docs-a-supercomputing-success>). For more than a year, a full-time NERSC postdoctoral fellow was assigned to the GTS project with the goal of improving its performance and scalability on the new NERSC petascale computer, Hopper, a 153,216-core Cray XE6 system running at 1.26 Pflops/sec peak. The project focused on the most communication-intensive part of the code, which is the particle “shift” step that follows the particle advance. Because of the toroidal domain decomposition, ions and electrons transiting around the torus change domain every few steps, depending on their velocity. In a single time step, about 10 percent of the particles move out of their current domain to move to the next in either direction. For a billion particles, that corresponds to more than 10 Gbytes of data being transmitted over the network at the same time.

To improve scalability, a new shift algorithm was developed using OpenMP task parallelism and one-sided communication calls implemented with Fortran co-arrays. In the original algorithm, the work was divided in two steps: a particle sort to identify and place in a buffer the particles that need to move out, and a communication step to

send those particles to their destinations. Instead of splitting the sorting of the particles and the communication step, the new scheme overlaps the two steps by filling a small buffer and issuing multiple small remote memory data transfers while the particles are being sorted. The one-sided communications don't involve the remote CPU due to the global address space support of the Cray XE6 Gemini interconnect. On that system, the new algorithm improved the shift step by more than 50 percent.³⁷ The same algorithm can be used by any PIC code using domain decomposition as one of its parallel schemes.

GTS continues to be improved by adding new physics and algorithms that lead to scientific discovery. For example, the recent addition of neoclassical physics, which is calculated self-consistently along with the turbulence physics, has led to new insights on the self-generation of noninductive currents in magnetically confined plasmas. These currents are of critical importance for long-pulse tokamak fusion experiments, such as ITER, as they're required to maintain high temperature and stability for the duration of the discharge. New GTS simulations carried out on NERSC's latest system, Edison, show that noninductive current generation is significantly enhanced in the presence of turbulent fluctuations triggered by an instability called *collision-less trapped electron mode* (CTEM). Unlike the nonturbulent version of this current, which is primarily carried by passing particles, magnetically trapped electrons carry the new current. The new understanding enabled by these simulations provides yet another way to control and enhance the confinement, and hence fusion production, in tokamak devices.

Tokamak Edge Simulations

In 2005, a new gyrokinetic PIC project was started by Choon-Seock Chang, then at the Courant Institute, New York University (NYU). His area of research is the often-neglected plasma edge region of tokamaks, which treats the plasma close to the wall of the device. This is a very complex region where plasma temperatures and densities drop dramatically, magnetic field lines stop forming well-defined closed surfaces, the plasma interacts with the wall components, neutral particles are an important part of the physics, and the effects of collisions become unavoidable. Furthermore, the useful delta-f multiscale method, which reduces sampling noise in particle simulations, isn't applicable to the edge region because the length scale of the plasma

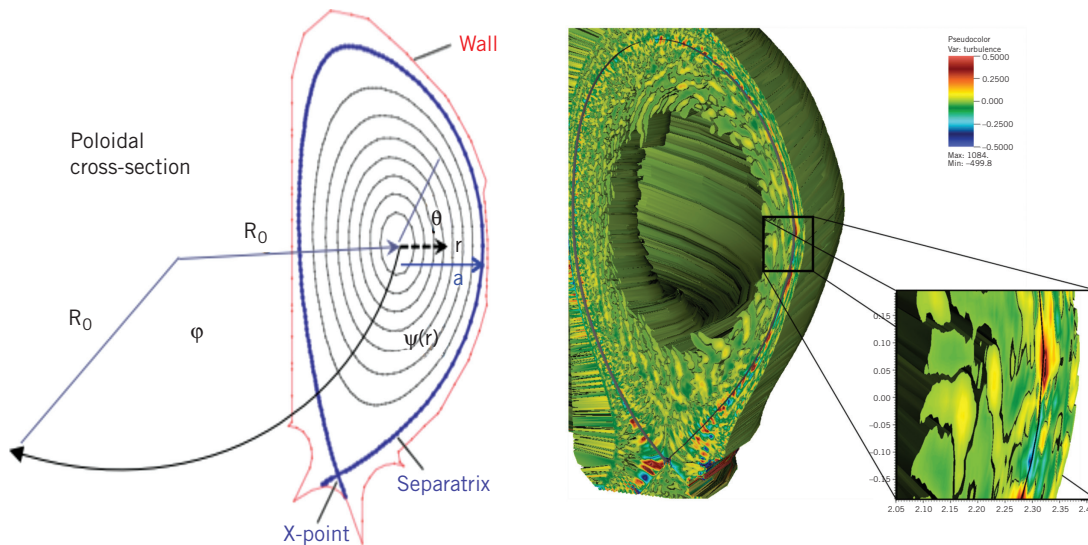


Figure 3. Domain simulated by the XGC1 code, which treats the plasma all the way to the tokamak wall. This capability allows for the study of the complex “edge” plasma region (enlarged section above).

gradients is of the same order as the length scale of the turbulence. The implementation of particle sources and sinks at the wall is also problematic in the delta-f framework.

Chang and postdoctoral fellow Seung-Hoe Ku undertook the development of XGC1, a full distribution function (full-f), global, gyrokinetic PIC code for the study of microturbulence in the edge region of tokamaks, including the interaction with the wall (Figure 3). As GTC did, XGC1 grew out of the Hamiltonian guiding center orbit algorithm of PPPL’s Orbit code, initially as a turbulence-free guiding-center neoclassical PIC code XGC0. With the establishment of the Prototype Fusion Simulation Project Center for Plasma Edge Simulation (CPES) in 2005, a project funded through the DoE program for Science Discovery through Advanced Computing (SciDAC), turbulence capability was added in XGC0, and the full-f 5D gyrokinetic code XGC1 was born.³⁸ The full-f kinetic description of the 5D plasma requires about 100 times more simulation particles than the delta-f description, but it allows for the dynamics and evolution of the plasma background.

Due to the complex geometry of the edge, and the region of opened magnetic field lines that it includes, it isn’t possible to use a field-line-following mesh everywhere and to carry out all operations in magnetic coordinates. Unlike in other gyrokinetic codes, the Lagrangian particle time advance is carried out in a cylindrical coordinate system,

while the Eulerian field solving is achieved on approximately field-aligned mesh, meaning that the mesh is mainly field aligned except close to the magnetic X-point where the mesh becomes too dense and anisotropic. An intricate unstructured mesh of triangular elements with varying resolution describes all the grid quantities in the simulation volume, which is extended to cover the interior of the whole tokamak to avoid inner-boundary effects.

It soon became clear that the extensive multiscale physics included in this code would make it effective as the core of a comprehensive tokamak simulation tool. Neutral particles recycled at the plasma wall are simulated together in a Monte Carlo method, with charge exchange and ionization interactions with the plasma. Plasma heating and torque sources are added to achieve steady background and turbulence profiles in a self-organized state. Nonlinear Fokker-Planck Coulomb operation is performed to describe the collisional process of the non-Maxwellian edge plasma.³⁹ As a result, XGC1 is able to study multiscale edge physics in realistic diverted magnetic field geometry, a unique capability among current gyrokinetic codes.

Initial development of XGC1, and its predecessor XGC0, was mostly done on NERSC’s Seaborg and Hopper. The first XGC0 neoclassical kinetic simulation in diverted magnetic geometry was performed on Seaborg, leading to the discovery

In general, our expectation is that improved methods and algorithms developed in one code should be deployable with some associated dedicated effort in the other advanced PIC codes.

of orbit loss dynamics during edge pedestal formation.⁴⁰ The first XGC1 gyrokinetic simulation of edge pedestal plasma in diverted DIII-D edge plasma was performed on Hopper, and yielded nonlinear, nonlocal ITG turbulence on the edge pedestal.⁴¹ Residual momentum transport mechanisms have been positively identified by injecting torque into the full- f 5D plasma to keep the toroidal rotation at zero everywhere.⁴²

The XGC1 code's mission moved forward on the new SciDAC-2 Center for Edge Physics Simulation (EPSI) in 2012. The applied mathematicians, the computer scientists, and the performance engineers in CPES and EPSI made XGC1 achieve linear scale up to the maximum number of processors on all US high-performance computers, including Hopper and Edison at NERSC, Titan at the Oak Ridge Leadership Computing Facility (OLCF), Mira at the Argonne Leadership Computing Facility (ALCF), and Blue Waters at National Center for Supercomputing Applications (NCSA). On Titan, XGC1 performs truly heterogeneous computing, sharing compute loads between CPUs and GPUs. Because XGC1 contains gyrokinetic physics of the whole plasma, unlike the perturbative plasma in delta- f codes, it requires many more computer resources than a delta- f code does. Consequently, XGC1 often runs on full systems at OLCF, ALCF, and NERSC.

Several scientific discoveries have been made recently on Edison—for example, a new “turbulent neoclassical” momentum transport in the edge region has been discovered that can explain the source of momentum in the core plasma.⁴³ Operating the ITER plasma with its toroidal rotation speed at a non-negligible fraction of the speed of sound could be an important ingredient in achieving its fusion goal through improved stability and transport. Unlike in the present tokamak plasmas, external momentum input into ITER plasma is difficult. Fortunately, the existing tokamak devices have found that a so-called “high-confinement” mode regime in the plasma can spontaneously develop a substantial amount of co-current directional toroidal rotation, with the momentum propagating rapidly into the core from the edge pedestal region.

Magnetic confinement fusion brings about very complex and diverse phenomena, and there is still much to learn and discover before achieving the goal of a turnkey fusion power plant. Computer simulation has been an invaluable tool in the quest to better understand and design fusion devices, and NERSC has been, and continues to be, a critical partner in our research progress. The same can certainly be said for all the other scientific projects that have used NERSC over the years. With each new system installed at NERSC, the increased performance and concurrency of these machines give scientists opportunities to push their applications further by adding more physics and increasing scalability. This isn't an easy task, which is why the NERSC staff, with its combined knowledge and expertise in high-performance computing, has been a tremendous help to all the projects.

Looking into the near future, NERSC will undergo a major change in its computer architecture. With the arrival of the next system, Cori, NERSC will move to the latest Intel Xeon Phi processor, a clear departure from the traditional superscalar processor that we've been using for more than 20 years. This accelerator-type architecture requires fine-grain parallelism with substantial multithreading to maximize performance. Fortunately, several of our global gyrokinetic codes currently include accelerator versions of the fine-grain multithreaded sections and continue to be improved. In addition to running at NERSC, the GPU version of XGC1 runs routinely on the hybrid CPU-GPU Cray XK7 Titan at OLCF. The GTC-Princeton (GTC-P) code, which was developed in collaboration with the Future Technologies group at LBNL, is our high-performance, high-scalability testbed code for new architectures. It has demonstrated delivery of scientific discovery advances enabled by computing at the extreme scale,^{44,45} can run efficiently on both the GPU and Xeon Phi systems, and is currently operating on the top seven supercomputers worldwide.

In general, our expectation is that improved methods and algorithms developed in one code should be deployable with some associated dedicated effort in the other advanced PIC codes. With Cori, we look forward to a new round of optimizations

and modifications for our codes, as always, with the help of NERSC's highly knowledgeable staff. The new NERSC Exascale Science Applications Program, or NESAP, has already selected XGC1 as one of 20 codes to receive dedicated help, and has given early access to all our other gyrokinetic PIC codes. This advance will let us continue to push the limits of our codes to achieve new scientific discoveries that will help us reach our ultimate goal of practical and affordable fusion energy for all. ■

Acknowledgments

Research at the Princeton Plasma Physics Laboratory is supported by the US Department of Energy, contract DE-AC02-09CH11466. Z. Lin is funded in part by the US Department of Energy (DOE) SciDAC GSEP Center. The research presented used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science, the US Department of Energy, under contract number DE-AC02-05CH11231.

References

1. J.M. Dawson, "Particle Simulation of Plasmas," *Rev. Modern Physics*, vol. 55, 1983, pp. 403–447.
2. A.B. Langdon and C.K. Birdsall, "Theory of Plasma Simulation Using Finite-Size Particles," *Physics of Fluids*, vol. 13, 1970, pp. 2115–2122.
3. H. Okuda and J.M. Dawson, "Theory and Numerical Simulation on Plasma Diffusion across a Magnetic Field," *Physics of Fluids*, vol. 16, 1973, pp. 408–426.
4. V.K. Decyk and J.M. Dawson, "Computer Model for Bounded Plasma," *J. Computational Physics*, vol. 30, 1979, pp. 407–427.
5. A.B. Langdon, "Energy-Conserving Plasma Simulation Algorithms," *J. Computational Physics*, vol. 12, 1973, pp. 247–268.
6. R.L. Morse and C.W. Nielson, "Numerical Simulation of Warm 2-Beam Plasma," *Physics of Fluids*, vol. 12, 1969, pp. 2418–2425.
7. O. Buneman, "Time-Reversible Difference Procedures," *J. Computational Physics*, vol. 1, 1967, pp. 517–535.
8. J. Denavit, "Numerical Simulation of Plasmas with Periodic Smoothing in Phase Space," *J. Computational Physics*, vol. 9, 1972, pp. 75–98.
9. J.P. Boris and K.V. Roberts, "The Optimization of Particle Calculations in 2 and 3 Dimensions," *J. Computational Physics*, vol. 4, 1969, pp. 552–571.
10. R.W. Hockney, S.P. Goel, and J.W. Eastwood, "Quiet High-Resolution Computer Models of a Plasma," *J. Computational Physics*, vol. 14, 1974, pp. 148–158.
11. C.K. Birdsall, and A.B. Langdon, *Plasma Physics via Computer Simulation*, McGraw-Hill, 1985.
12. R.W. Hockney and J.W. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill, 1981.
13. W.W. Lee, "Gyrokinetic Approach in Particle Simulation," *Physics of Fluids*, vol. 26, 1983, pp. 556–562.
14. W.W. Lee, "Gyrokinetic Particle Simulation Model," *J. Computational Physics*, vol. 72, 1987, pp. 243–269.
15. S.E. Parker and W.W. Lee, "A Fully Nonlinear Characteristic Method for Gyrokinetic Simulation," *Physics of Fluids B*, vol. 5, no. 1, 1993, pp. 77–86.
16. S.E. Parker, W.W. Lee, and R.A. Santoro, "Gyrokinetic Simulation of Ion Temperature Gradient Driven Turbulence in 3D Toroidal Geometry," *Physical Rev. Letters*, vol. 71, no. 13, 1993, pp. 2042–2045.
17. R.D. Sydora, "Toroidal Gyrokinetic Particle Simulations of Core Fluctuations and Transport," *Physica Scripta*, vol. 52, 1995, pp. 474–480.
18. R.D. Sydora, V.K. Decyk, and J.M. Dawson, "Fluctuation-Induced Heat Transport Results from a Large Global 3D Toroidal Particle Simulation Model," *Plasma Physics and Controlled Fusion*, vol. 38, 1996 pp. A281–A294.
19. J.-N. Leboeuf et al., "Effect of Externally Imposed and Self-Generated Flows on Turbulence and Magnetohydrodynamic Activity in Tokamak Plasmas," *Physics Plasmas*, vol. 7, 2000, pp. 1795–1801.
20. M.J. LeBrun et al., "Toroidal Effects on Drift Wave Turbulence," *Physics of Fluids B*, vol. 5, 1993, pp. 752–773.
21. R.B. White and M.S. Chance, "Hamiltonian Guiding Center Drift Orbit Calculation for Plasmas of Arbitrary Cross Section," *Physics Fluids*, vol. 27, no. 10, 1984, pp. 2455–2467.
22. R.B. White and A.H. Boozer, "Rapid Guiding Center Calculations," *Physics Plasmas*, vol. 2, no. 8, 1995, pp. 2915–2919.
23. Z. Lin and W.W. Lee, "Method for Solving the Gyrokinetic Poisson Equation in General Geometry," *Physical Rev. E*, vol. 52, no. 5, 1995, pp. 5646–5652.
24. Z. Lin et al., "Turbulent Transport Reduction by Zonal Flows: Massively Parallel Simulations," *Science*, vol. 281, 18 Sept. 1998, pp. 1835–1837.
25. Z. Lin et al., "Size Scaling of Turbulent Transport in Magnetically Confined Plasmas," *Physical Rev. Letters*, vol. 88, no. 19, 2002, article 195004-1.
26. S. Ethier et al., "Large-Scale Gyrokinetic Particle Simulation of Microturbulence in Magnetically Confined Plasmas," *IBM J. Research and Development*, vol. 52, 2008, pp. 105–115.
27. W. Zhang, Z. Lin, and L. Chen, "Transport of Energetic Particles by Microturbulence in Magnetized

- Plasmas," *Physical Rev. Letters*, vol. 101, 2008, article 095001.
28. Y. Yong Xiao and Z. Lin, "Turbulent Transport of Trapped Electron Modes in Collisionless Plasmas," *Physical Rev. Letters*, vol. 103, 2009, article 085004.
 29. H.S. Zhang, Z. Lin, and I. Holod, "Nonlinear Frequency Oscillation of Alfvén Eigenmodes in Fusion Plasmas," *Physical Rev. Letters*, vol. 109, 2012, article 025001.
 30. Z. Wang et al., "Radial Localization of Toroidicity-Induced Alfvén Eigenmodes," *Physical Rev. Letters*, vol. 111, 2013, article 145003.
 31. K. Madduri et al., "Gyrokinetic Particle-in-Cell Optimization on Emerging Multi- and Manycore Platforms," *Proc. Int'l Conf. High Performance Computing, Networking, Storage and Analysis*, 2011; <http://dl.acm.org/citation.cfm?id=2063415>.
 32. W.X. Wang et al., "Gyro-Kinetic Simulation of Global Turbulent Transport Properties in Tokamak Experiments," *Physics of Plasmas*, vol. 13, no. 9, 2006, article 092505.
 33. W.X. Wang et al., "Nonlinear Flow Generation by Electrostatic Turbulence in Tokamaks," *Physics Plasmas*, vol. 17, 2010, article 072511.
 34. S. Ethier et al., "Global Gyrokinetic Simulation of Electron Temperature Gradient Turbulence and Transport in NSTX Plasmas," *Proc. 23rd IAEA Fusion Energy Conf.*, 2010; www.naweb.iaea.org/napcl/physics/fec/fec2010/html/node326.htm.
 35. W.X. Wang et al., "Trapped Electron Mode Turbulence Driven Intrinsic Rotation in Tokamak Plasmas," *Physical Rev. Letters*, vol. 106, 2011, article 085001.
 36. W.X. Wang et al., "Gyrokinetic Studies on Turbulence-Driven and Neoclassical Nondiffusive Toroidal-Momentum Transport and the Effect of Residual Fluctuations in Strong ExB Shear," *Physical Rev. Letters*, vol. 102, 2009, article 035005.
 37. R. Preissl et al., "Multithreaded Global Address Space Communication Techniques for Gyrokinetic Fusion Applications on Ultra-Scale Platforms," *SC'11 Proc. Int'l Conf. High Performance Computing, Networking, Storage and Analysis*, 2011, article 78; <http://dx.doi.org/10.1145/2063384.2071033>.
 38. S. Ku, C.S. Chang, and P.H. Diamond, "Full-f Gyrokinetic Particle Simulation of Centrally Heated Global ITG Turbulence from Magnetic Axis to Edge Pedestal Top in a Realistic Tokamak Geometry," *Nuclear Fusion*, vol. 49, 2009, article 115021.
 39. E.S. Yoon and C.S. Chang, "A Fokker-Planck-Landau Collision Equation Solver on Two-Dimensional Velocity Grid and Its Application to Particle-in-Cell Simulation," *Physics Plasmas*, vol. 21, 2014, article 032503.
 40. C.S. Chang, S. Ku, and H. Weitzner, "Numerical Study of Neoclassical Plasma Pedestal in a Tokamak Geometry," *Physics Plasmas*, vol. 11, 2004, article 2649.
 41. C.S. Chang et al., "Compressed Ion Temperature Gradient Turbulence in Diverted Tokamak Edge," *Physics Plasmas*, vol. 16, 2009, article 056108.
 42. S. Ku et al., "Physics of Intrinsic Rotation in Flux-Driven ITG Turbulence," *Nuclear Fusion*, vol. 52, 2012, article 063013.
 43. J. Seo et al., "Intrinsic Momentum Generation by a Combined Neoclassical and Turbulence Mechanism in Diverted DIII-D Plasma Edge," *Physics Plasmas*, vol. 21, 2014, article 092501.
 44. B. Wang et al., "Kinetic Turbulence Simulations at Extreme Scale on Leadership-Class Systems," *SC'13 Proc. Int'l Conf. High Performance Computing, Networking, Storage and Analysis*, 2013, article 82; <http://dx.doi.org/10.1145/2503210.2503258>.
 45. W. Tang, B. Wang, and S. Ethier, "Scientific Discovery in Fusion Plasma Turbulence Simulations at Extreme Scale," *Computing in Science & Eng.*, vol. 16, no. 5, 2014, pp. 44–52.
-
- Stephane Ethier** is deputy head of the Computational Plasma Physics Group at the US Department of Energy's Princeton Plasma Physics Laboratory. His research interests include high-performance computing and large-scale particle simulations of micro-turbulence in magnetic confinement fusion devices. Ethier has a PhD in physics from the Institut National de la Recherche Scientifique: Energie et Materiaux in Montreal, Canada. He's a member of the American Physical Society (APS) and ACM. Contact him at ethier@pppl.gov.
-
- Choon-Seock Chang** is a principal research physicist at the Princeton Plasma Physics Laboratory and head of the multi-institutional multidisciplinary US SciDAC Center for Edge Physics Simulation (EPSI), awarded jointly by US DoE's Office of Fusion Energy Science and Office of Advanced Scientific Computing Research. Chang has a PhD in physics from the University of Texas at Austin. He's a fellow of the APS. Contact him at cschang@pppl.gov.
-
- Seung-Hoe Ku** is research physicist at the US Department of Energy's Princeton Plasma Physics Laboratory and lead author of XGC1, X-point included gyrokinetic code in the SciDAC-3 Center for Edge Physics Simulation (EPSi). Ku has a PhD in physics from the Korea Institute of Science and Technology. He's a member of the APS. Contact him at sku@pppl.gov.

Wei-li Lee is a Distinguished Laboratory Research Fellow at Princeton Plasma Physics Laboratory. His main area of research is the nonlinear gyrokinetic theory and simulation of microturbulence, a field that he pioneered in the mid-1980s. Contact him at wwlee@pppl.gov.

Weixing Wang is a principal research physicist at the US Department of Energy's Princeton Plasma Physics Laboratory. His research areas and interests include theory and computation of plasma micro-instabilities, turbulent and collisional transport in magnetic fusion experiments, gyrokinetic simulation, and advanced simulation algorithms. Wang is a member of the APS. Contact him at wwang@pppl.gov.

Zhihong Lin is a professor of physics and director of DOE's SciDAC GSEP Center at the University of California, Irvine. His research interests include large-scale simulations of waves, instabilities, and turbulence in magnetized plasmas. Lin has a PhD in plasma physics from Princeton University. He's a fellow of the APS. Contact him at zhihongl@uci.edu and <http://phoenix.ps.uci.edu/zlin>.

William Tang is a principal research physicist at the Princeton Plasma Physics Laboratory at Princeton University, where he's also a lecturer/professor in the Department of Astrophysical Sciences and a member of the Executive Board of the Princeton Institute for Computational Science and Engineering (PICSciE), which he helped establish. His research interests include mathematical formalism and computational applications dealing with electromagnetic kinetic plasma behavior in complex geometries; simulation of complex transport dynamics driven by microturbulence in plasmas; and validated predictive integrated modeling in tokamak plasmas. Tang has a PhD in physics from the University of California, Davis. He's a fellow of the APS and recipient of the Chinese Institute of Engineers-USA Distinguished Achievement Award. Contact him at wtang@princeton.edu.



Selected articles and columns from IEEE Computer Society publications are also available for free at <http://ComputingNow.computer.org>.

ADVERTISER INFORMATION

Advertising Personnel

Marian Anderson: Sr. Advertising Coordinator
Email: manderson@computer.org
Phone: +1 714 816 2139 | Fax: +1 714 821 4010

Sandy Brown: Sr. Business Development Mgr.
Email: sbrown@computer.org
Phone: +1 714 816 2144 | Fax: +1 714 821 4010

Advertising Sales Representatives (display)

Central, Northwest, Far East:
Eric Kincaid
Email: e.kincaid@computer.org
Phone: +1 214 673 3742
Fax: +1 888 886 8599

Northeast, Midwest, Europe, Middle East:
Ann & David Schissler
Email: a.schissler@computer.org, d.schissler@computer.org
Phone: +1 508 394 4026
Fax: +1 508 394 1707

Southwest, California:
Mike Hughes
Email: mikehughes@computer.org
Phone: +1 805 529 6790

Southeast:
Heather Buonadies
Email: h.buonadies@computer.org
Phone: +1 973 304 4123
Fax: +1 973 585 7071

Advertising Sales Representatives (Classified Line)

Heather Buonadies
Email: h.buonadies@computer.org
Phone: +1 973 304 4123
Fax: +1 973 585 7071

Advertising Sales Representatives (Jobs Board)

Heather Buonadies
Email: h.buonadies@computer.org
Phone: +1 973 304 4123
Fax: +1 973 585 7071