ACCELERATING NANOSCALE TRANSISTOR INNOVATION WITH NEMO5

Allocation: NSF/2.48 Mnh
PI: Gerhard Klimeck[1]
Collaborators: Jim Fonseca[1]; Tillmann Kubis[1]; Daniel Mejia[1]; Bozidar Novakovic[1]; Michael Povolotskyi[1]; Mehdi Salmani-Jelodar[1]; Harshad Sahasrabudhe[1]; Evan Wilson[1]; Kwok Ng[2]

Intel Corporation; PETSc Development Team, Argonne National Laboratory; NVIDIA Corporation

[1]Purdue University
[2]Semiconductor Research Corporation

SCIENTIFIC GOALS

According to the International Technology Roadmap for Semiconductors (ITRS), transistor size will continue to decrease in the next 10 years, but foundational issues with current technology means new approaches must be pursued. The number of atoms in critical dimensions is now countable. As the materials and designs become more dependent on atomic details, the overall geometry constitutes a new material that cannot be found as such in nature.

We at the Institute for Nanoelectronic Modeling (iNEMO) have used Blue Waters to drive a variety of scientific projects related to semiconductor device modeling with atomistic detail. The NEMO software tool suite provides input to ITRS and is also used by leading semiconductor firms to design future devices. The current version of the software, NEMO5, is developed and used by iNEMO to address fundamental questions on a variety of semiconductor devices.

iNEMO research on Blue Water encompasses scientific work including stacked quantum dots, disordered transistor leads, new ITRS projections, and materials such as bilayer graphene and topological insulators. Key aspects of these research projects would not be possible without Blue Waters. Future goals include more advanced models such as electron-electron scattering and time dependent simulations.

ACCOMPLISHMENTS TO DATE

ITRS. We used the NEMO software suite on Blue Waters to calculate design parameters for future devices and these results have been included in the 2013 ITRS. Simulations found important deviations in the characteristics of devices as they are scaled down and raised questions about future device design.

Quantum Dot Donor Configuration Interaction. Our simulations, in close collaboration with leading experimental groups, help to understand and design 2-qubit gates for semiconductor quantum computing. Blue Waters simulations have elucidated charging energies and the configurations of the two-electron ground state of a single negative donor as a function of depth and external electric fields. Two-electron ground states of donor molecules match well to experimental results and a new quantum dot design has been proposed based on the electric field dependency of the exchange coupling.

Topological Insulators. Surface states of topological insulator nanowires are expected to serve as scattering-free charge conductors at room temperature. Devices made of topological insulators have been simulated using the tight binding method, which treats each atom individually. This approach, in conjunction with results from our experimental collaborators, shows that surface state dispersion is geometry dependent, a result that cannot be described by simpler approaches. Because of the size of these devices and the necessary atomic representation, it would not be possible to simulate them without a machine of Blue Waters' caliber.

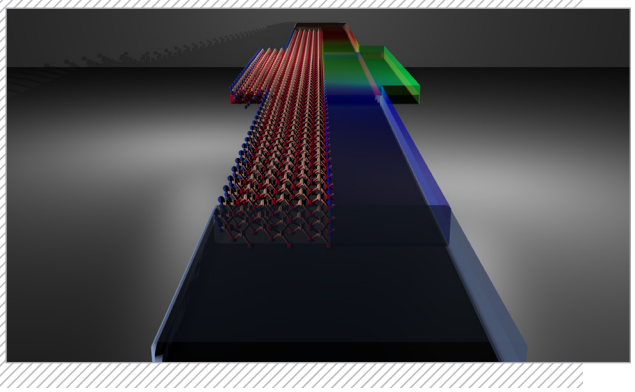

Figure 1. Potential along an InAs Ultra-Thin Body (UTB) transistor.

Disorder in Semiconductor Device Leads. Research into non-ideal device leads has focused on alloy disorder and how its tuning can optimize heat flow in ITRS devices to avoid Joule heating and recycle heat into usable electric power.

Quantum Dot Stacks. Work with the Air Force Research Lab led to NEMO5 development for sub-monolayer quantum dot stacks to allow tuning of optical emission and absorption of extended emitter and detector systems in a wide range of wavelengths. An optimized strain model that includes the effect of Coulomb interaction was developed and shows that modeling can accurately determine absorption spectra of these experimental devices.

Figure 2. Atomistic representation of Si, represented as diamond structure crystals.

Bilayer Graphene Transistors. Graphene promises improved device characteristics if a suitable bandgap can be opened. Recent studies with NEMO5 have simulated transport in bilayer graphene with a tight-binding approach and dynamically controlled bandgap.

GPU specific development. NEMO5 has achieved efficient scalability past 100 nodes on Blue Waters and performance that shows a single NVIDIA Kepler K20x GPU can provide as much processing power as 40 AMD Bulldozer cores. Additionally, we have developed an interface between Argonne National Lab's PETSc library for scientific computing and MAGMA, a library for linear algebra on GPUs. This interface is schedule to be released to the scientific community as part of a PETSc release.

HOW BLUE WATERS PROJECT STAFF HELPED

Blue Waters staff has been helpful in all regards. Examples include: monthly user webcasts with system updates and training presentations, custom scripts to analyze our usage, assistance building specialized numerical libraries, testing NEMO5 code for compatibility with forthcoming programming environment updates, and dealing with the occasional lost RSA token. The yearly PRAC and advanced user workshops have been especially beneficial for graduate students learning about HPC.

WHY THE RESEARCH MATTERS

The U.S. semiconductor industry is one of the largest export industries. The global semiconductor device market is over $300 billion and the U.S. holds more than one third of this market. The U.S. is a market leader and produces a significant number of high-paying, high-technology jobs. At the same time, the end of Moore's law scaling as we know it will be reached in ten years with device dimensions expected to be about 5 nm long and 1 nm, or about 5 atoms, in its critical active region width. Further improvements in these dimensions will come only through detailed and optimized device design and better integration.

WHY BLUE WATERS

The simulations required to conduct our research simply cannot be accomplished in a timely manner on less powerful machines. Without Blue Waters, iNEMO research is, at best, hampered, and at worst, impossible. For instance, a simplistic transport calculation of a 50 nm long wire with a 3 nm diameter (about 28 nm^2 in cross section) requires around 1 Tflop/s for a single energy point using a non-equilibrium Green's Function. Resolution of a device's characteristics requires about 1,000 energy points, and this calculation must be repeated perhaps a dozen times for a full current-voltage sweep. Even with the recursive Green's function, the computational time scales with the cube of cross sectional area (relative to the direction of the electron flow) and linearly with the length of the device. The treatment of a technically currently relevant finFET device would require an atomistic resolution of a device with a cross-section around (20x40) nm^2, which includes the core semiconductor and the surrounding gate material. Blue Waters

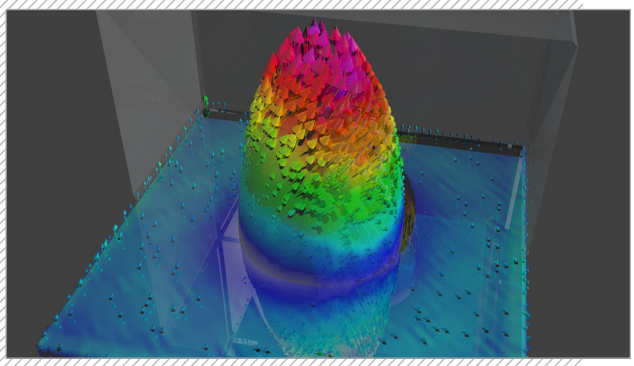

Figure 3. InAs-GaAs quantum dot strain displacement.

allows us to get results in a day, rather than weeks on other computer systems. Additionally, some devices, such as those in the topological insulator work, are simply too physically large (due to the number of atoms and size of related tight-binding matrices) to fit on less powerful systems.

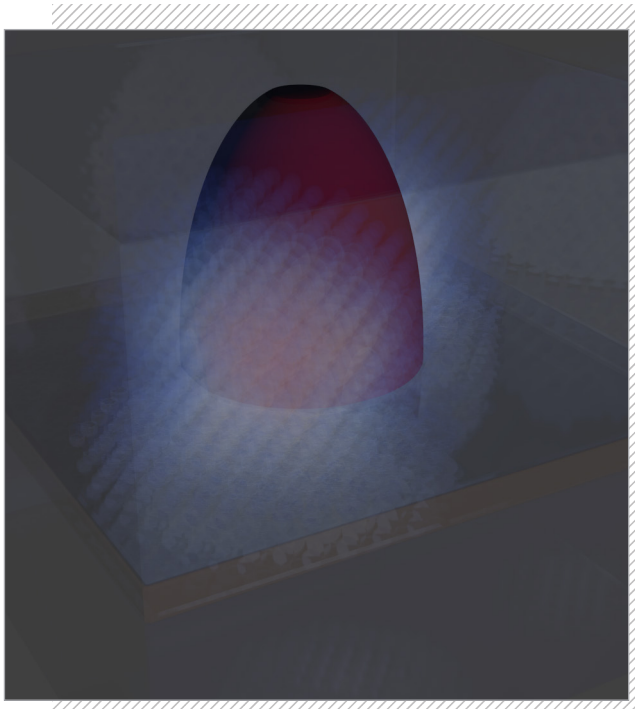

Figure 4. InAs-GaAs quantum dot stationary wave functions

PRE-PETASCALE PREPARATION

Codes in the NEMO tool suite have been shown to scale well on leadership-class machines. A previous version of NEMO5, OMEN, demonstrated almost perfect scaling to 222,720 cores and 1.44 Pflop/s, the first engineering code to deliver a sustained 1.4 Pflop/s on over 220,000 cores on Jaguar. NEMO5 is a more general code but implements the same underlying numerical approaches and framework as OMEN. Scalable algorithms that primarily use dense/sparse-dense matrix-matrix multiplication and matrix inversion have been used to take advantage of Blue Waters' GPU capabilities. We did profiling and benchmarking work on NEMO5's infrastructure to vastly improve scaling into the 1,000 XE node realm for the simulation studies described previously.

LOOKING FORWARD TO THE NEXT TRACK-1 SYSTEM

One-hundred-million atom 3D simulations for many crystals using Spin and/or Classical Multi-Physics, full transport, and problem-specific assumptions would be feasible on future Track-1 systems. A Track-1 system would allow ultra-scaled transistors to be simulated with important physical phenomena such as scattering, alloy disorder, and time dependence, as well as new device materials and designs. Multiscale simulations would allow for inclusion of not only the transistor part of the device, but also the devices' environment, such as the metal leads—which have an increasingly strong effect on device performance as the channel region is diminished. Statistical ensembles would also be possible to investigate process and fabrication effects on physical models and device characteristics.

COMMUNITY IMPACT

NEMO5 source code is available free for academic use with an associated nanoHUB.org user support group where users can interact, get support from the development team, access a variety of tutorials and even simulate devices with precompiled NEMO5 binaries on Purdue's HPC resources. The aforementioned PETSc-MAGMA interface is undergoing testing and will be released by Argonne to the greater scientific community.

PUBLICATIONS

Salmani-Jelodar, M., J. D. Bermeol, S. Kim, and G. Klimeck, ITRS Tool on NanoHUB. *NanoHUB User Conf.*, Phoenix, Ariz., April 9-11, 2014.

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications.