EARLY EXA-SCALE SCIENCE OPPORTUNITY: EDGE PLASMA PHYSICS

MARK F. ADAMS

Prepared for the DOE Workshop on Integrated Simulations for Magnetic Fusion Energy Sciences

Section (D) on Multiphysics and multiscale coupling

Introduction: Particle in cell (PIC) methods are an effective and popular approach to discretizing high dimensional PDEs such as magnetized plasmas. For example, the premier extreme-scale tokamak burning edge plasma physics code in the world, XGC in the Edge Plasma Simulation Institute (EPSI), is one of the DOE's OFES flagship projects. PIC methods use particles to discretize high dimensional systems such as a 5D gyrokinetic Vlasov equation, and use grids to discretize low dimensional systems such as a 3D Poisson equation. Particle processes are highly concurrent, use regular array data structures, and are amenable to distributed memory processing, all of which have contributed to the success of PIC methods in extreme-scale high dimensional simulations. At a high level, PIC methods have as extended low dimensional field methods with a high dimensional phase space component to resolve multiscale phenomenon with tight coupling between scales and physics. Optimal Poisson solvers are multilevel in and of themselves; PIC methods extend equation like Maxwell's to higher resolution phase space equations like Vlasov and are, thus, a type of multilevel method. Early science from exa-scale compute platforms will likely come from extreme multiscale science simulations, which require multilevel methods for efficient computational complexity. Because they are able to resolve large ranges of scales efficiently, PIC methods are ideal discretization techniques for early exa-scale science applications.

In addition to extreme multiscale, many science applications have a distinguished coordinate (DC) with associated anisotropic dynamics. For example, the physics of tokamak plasmas is dominated by a strong magnetic field, which is used to contain the burning plasma. PIC methods split the fields with respect to this DC and often use aligned grids to ensure, among other things, that parallel derivatives are computed without numerical transverse diffusion errors.

Many PIC codes have flourished in extreme-scale computing with replicated field data because, for example, PIC methods for non-Maxwellian physics use ∼10,000 particles per cell and the grids are commensurately small. Distributed grids complicate PIC codes significantly, but are required for future extreme-scale simulations.

The opportunity: PIC methods are an ideal discretization technique for early exa-scale science. There are many challenges in understanding the modeling of plasmas, which are being actively developed in the EPSI project, among many others, but the underlying numerical methods, mesh and particle management, and solvers are in need of a significant refactoring; this technology can be developed in parallel with current physics research if ASCR researchers and OFES computational physicists work together from the beginning.

The challenges: First among the needs of future extreme-scale tokamak PIC codes is distributed mesh and particle management. In the last several years, great strides have been made in distributed grid management with semi-structured methods that are ideally suited to tokamak plasma applications. These methods use a hybrid or multilevel approach in which an unstructured

---

Scalable Solvers Group, Lawrence Berkeley National Laboratory, Berkeley, CA, `mfadams@lbl.gov` 609.206.2040.

coarse grid is redundantly stored and each cell in this coarse grid is refined with a regular tree data structure to form a forest of octrees (in 3D). These methods have matured to the point that they are availably in a library, p4est, and being developed in PETSc.

Tree-based approaches are ideally suited to high performance methods on future architectures for several converging reasons. The kernel computations are amenable to the high number of vector lanes characteristic of current and future architectures. Structured grids facilitate mathematically fast geometric multigrid solvers, which minimize latency and data movement costs. This hybrid of tree-based structured grids with an unstructured coarse grid strikes an ideal balance between meshing flexibility and kernel simplicity for many applications. These distributed grid data structures are efficient for many PIC processes such as the particles search for finding the cell for a particle, and they naturally enforce tight coupling of particle and grid data.

Distributed mesh and particle management invariably requires different grids than the Poisson solver because the particle work does not align well with efficient solver grid data layouts in general. For example the EPSI project is currently working on a dual mesh algorithms in which a stretched, field line following (twisting around the tokamak) non-conforming mesh is used for particle management and in processing the collision operator, and an axisymmetric, conforming finite element grid is used for the Poisson solver. The mesh-particle interactions such as charge deposition and electric field calculations require overlapping representations of the solver grid on each process. Moreover, the desire to maintain data locality between these two grids requires multi-objective and arbitrary metric space partitioning.

The traditional organization of PIC codes naturally reflects the mathematics and the algorithm: deposit charge, Poisson solve, compute gradients at particle positions, compute collisions, push particles, compute diagnostics, but does not map well to processing on modern architectures. This approach requires that particles be read (streamed) into registers several times in each time step or stage. This excess data movement was not a constraint in the past, but it is now. Data movement can be reduced significantly by reorganizing PIC code by streaming particles into registers and fusing particle processing. For example, push each particle, deposit its charge for Poisson solve (on a spacial grid), and deposit its charge for Fokker-Plank collision solve (on a velocity grid); and then solve the Fokker-Plank and Poisson systems at each time step or stage.

Impact: Fusion energy science is a core competence of DOE and is an integral part of the DOE mission. EPSI is the flagship simulation project for tokamak burning plasmas in DOE. Efficient distributed grid and particle data algorithms will allow extreme scale PIC codes to effectively utilize exa-scale compute resources and resolve physics unattainable with current data models. Addressing this problem is critical for the continued development of this component of the DOE science portfolio and magnetically confined plasma physics in general.

Future architectures will disrupt applications that wish to exploit these resources. Early users of extreme-scale or exa-scale resources will use algorithms that adapt most easily to these new architectures. As architectures and programming models mature, a broader range of applications will inevitably be accommodated, but achieving early success is critical for the health of high performance computing. PIC methods have a natural multilevel structure, the most efficient organization for scalable global algorithms, in which the majority of the work is local processing. The scientific demands of understanding fusion plasma physics and the ideal fit of PIC methods to modern architectures makes these applications ideal vehicles for early science on exa-scale platforms.