Recent results and future challenges for large scale Particle-In-Cell simulations of plasma-based accelerator concepts

Huang, C.

2009-10-12

HIFAN 1720

Recent results and future challenges for large scale Particle-In-Cell simulations of plasma-based accelerator concepts

by
C. Huang1, W. An1, V.K. Decyk1, W. Lu1, W.B. Mori1, F.S. Tsung1, M. Tzoufras2, S. Morshed3, T. Antomsen3, B. Feng4, T Katsouleas5, R.A. Fonseca6, S.F. Martins6, J. Vieira6, L.O. Silva6, C.G.R. Geddes7, E Cormier-Michel7, J.-L. Vay7, E. Esarey7, W.P. Leemans7, D.L. Bruhwiler8, B. Cowan8, J.R. Cary8,9, K. Paul8

1 UCLA
2 University of Oxford
3 University of Maryland
4 University of Southern CA
5 Duke University
6 Instituto Superior Tecnico
7 LBNL
8 Tech-X Corporation
9 University of Colorado

Accelerator Fusion Research Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

June 2009

This work was supported by the Director, Office of Science, Office of Fusion Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Recent results and future challenges for large scale Particle-In-Cell simulations of plasma-based accelerator concepts

C. Huang[1], W. An[1], V.K. Decyk[1], W. Lu[1], W.B. Mori[1], F. S. Tsung[1], M. Tzoufras[2], S. Morshed[3], T. Antonsen[3], B. Feng[4], T Katsouleas[5], R.A. Fonseca[6], S.F. Martins[6], J. Vieira[6], L.O Silva[6], E. Esarey[7], C. G. R. Geddes[7], W. P. Leemans[7], E. Cormier-Michel[7], J.-L. Vay[7], D. L. Bruhwiler[8], B. Cowan[8], J. R.Cary[8,9], K. Paul[8]

[1] University of California, Los Angeles, CA 90095, USA
[2] University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK
[3] University of Maryland, College Park, MD 20742, USA
[4] University of Southern California, Los Angeles, CA 90089, USA
[5] Duke University, Durham, NC 27708, USA
[6] Instituto Superior Tecnico, Lisbon 1049-001, Portugal
[7] Lawrence Berkeley National Lab, Berkeley, CA 94720, USA
[8] Tech-X Corporation, Boulder, CO 80303, USA
[9] University of Colorado, Boulder, CO 80309, USA

Email: huangck@ee.ucla.edu

Abstract. The concept and designs of plasma-based advanced accelerators for high energy physics and photon science are modelled in the SciDAC COMPASS project with a suite of Particle-In-Cell codes and simulation techniques including the full electromagnetic model, the envelope model, the boosted frame approach and the quasi-static model. In this paper, we report the progress of the development of these models and techniques and present recent results achieved with large-scale parallel PIC simulations. The simulation needs for modelling the plasma-based advanced accelerator at the energy frontier is discussed and a path towards this goal is outlined.

1. Introduction

Particle accelerators have been the major tool for experimental high energy physics and photon science. The largest particle accelerator ever built to date, the Large Hadron Collider (LHC) at CERN, is based on the superconducting RF accelerator technology. It is 27 kilometers in length and cost 6 billion dollars. Similarly the world's first hard X-ray free electron laser, the Linac Coherent Light Source (LCLS), also relies on a kilometer-long conventional RF accelerator to accelerate an electron beam to 14.35 GeV for x-ray FEL production. A possible way to reduce the size and cost of next generation accelerators by an order of magnitude or more is to use relativistically moving plasma waves as the accelerating structure [1, 2]. Relativistic plasma waves can be excited by the space charge forces of a charged particle beam in the Plasma Wakefield Acceleration (PWFA) scheme or by the radiation pressure of a high intensity laser in the Laser Wakefield Acceleration (LWFA) scheme.

Recent experiments on plasma-based advanced accelerators have achieved 30~50 GeV/m acceleration gradient and produced GeV class electron beams with percents of energy spread from the

LWFA scheme over cm distances. In recent PWFA experiments similar gradients were demonstrated over meter distances and 80 GeV electrons were produced. Simulations show that properly phased electrons can be accelerated while keeping the energy spread near or below a percent using the PWFA scheme. However, the energy frontier for high energy physics is in the TeV range, while 1-20 GeV electron beams with small angular divergence are desirable for applications in photon science. The key challenge to use plasma-based accelerators for these applications is to extend the acceleration length over meters of plasma with stringent requirements on the beam parameters and quality. It is also a challenge to accurately simulate these next generation experiments.

The interactions among the plasma particles and with the driver in the self-consistent electromagnetic fields they produce are fully modeled with Particle-In-Cell (PIC) codes [3, 4]. These models provide the most fundamental, microscopic level description of the interactions through Maxwell's equations and Newton's laws; therefore, they are also very computation intensive. PIC models are differentiated by the kinds of fields retained in the model, the manner in which the fields are solved, and the charge density and current density are deposited.

Within the SciDAC COMPASS (Community Petascale Project for Accelerator Science and Simulation) project, a suit of parallel PIC codes (VORPAL, OSIRIS, WARP, QuickPIC) are actively maintained and enhanced to model particle beam generation, acceleration for the energy frontier and radiation sources for biological and material applications using the plasma-based accelerator concepts. VORPAL [5], OSIRIS [6], Warp [7], are state-of-the-art, fully explicit, multi-dimensional, fully parallelized, fully relativistic, PIC codes. These PIC codes are capable of reproducing the detailed physics at the expense of computation power. They have also been successfully benchmarked against each other and used to validate other reduced algorithms [8]. Recently the speed of these PIC codes has been greatly enhanced through a boosted frame algorithm by utilizing the fact that the range of space/time scale is not a Lorentz invariant [9]. This speed improvement enables 3D full scale simulations of near term LWFA experiments in great detail.

For the LWFA problem, the ponderomotive guiding center approximation can be used to remove the need of resolving the finest time scale of the laser oscillation. In the VORPAL code, this is implemented in the laser envelope model, which applies the approximation without loss of kinetic effects such as particle trapping [10]. Pulse depletion is also modeled accurately for long propagation distances, although the corresponding spectral broadening eventually breaks the assumptions of the algorithm. The model is therefore suitable to simulate both long propagation distances and injection mechanisms, with orders of magnitude speedup relative to the standard PIC algorithm, depending on application. The QuickPIC code is a highly efficient, fully parallelized, fully relativistic, three-dimensional PIC code for simulating LWFA/PWFA schemes [11] that utilizes the quasi-static approximation together with the ponderomotive guiding center approximation. Under the quasi-static approximation, the time scale of the evolution of the driver is separated out from the plasma evolution and a fully three-dimensional electromagnetic field solve and particle push is reduced to a sequence of transverse two-dimensional field solves and particle pushes. Overall, this algorithm speeds up the computational time by 2 to 4 orders of magnitude without losing accuracy for problems of interest.

Using this collection of PIC codes with peta-scale computers, we can carry out high resolution high fidelity full-scale simulations to determine if there are physics obstacles to a plasma based accelerator design before large capital expenditures are spent building key components. Currently, a 1-5 GeV single stage LWFA can be modeled in 3D using full PIC codes. Reduced models and techniques are being developed to allow simulation of multi-GeV stages, including the boosted frame algorithm, density scaled simulation, and the envelope or quasi-static models, and with careful benchmarking these are making it possible to simulate 10~50 GeV single stage LWFA/PWFA in 3D. In this paper, we describe recent large-scale kinetic simulations of plasma-based accelerators by the SciDAC COMPASS project. Section 2 describes recent progress and results from simulations to model future PW class LWFA experiments towards a Laser-Plasma Linear Collider (LPLC). Section 3 describes modeling of a TeV PWFA Linear Collider (PWFA-LC) concept using the quasi-static model. Section

4 provides an overview of the needs for modeling LWFA/PWFA for the energy frontier. Finally, the path towards peta-scale and beyond simulations of LWFA/PWFA is outlined.

2. Recent progress and results of simulation of LWFA

2.1. Modelling next generation LWFA experiments and LPLC designs

The simulation of the long distances associated with the next generation of lasers systems can be achieved with the use of reduced codes and methods including QuickPIC, VORPAL envelope, and also with the boosted frame approach. These techniques, combined with benchmarkings against full PIC simulations for short distances and scaled simulations, allow one to address the range of physics required for LPLC modules and other applications.

Three dimensional simulation with QuickPIC in the laboratory frame and OSIRIS in the boosted frame [12] were performed for different possible configurations for next generation experiments with a ~300J laser using weakly-nonlinear/nonlinear regimes and self-injected/externally-injected electron beams. QuickPIC results for self-injection and external-injection/external-guiding regimes are shown in Fig. 1. Self-injection (modeled with test particles) is obtained with a moderate intensity pulse (a_0=5.8) propagating in a 2.7×10^{17} cm^{-3} density plasma. The external-injection case uses a_0=2.0, and a plasma density of 2.2×10^{16} cm^{-3}. Results agree with theoretical scalings [13] for the accelerating gradients and the final electron beam energies. These scalings show how to scale results from one density to another for matched laser spot sizes. We also note the extremely stable propagation and wake excitation in the external injection scenario was seen for the meter scale distance.

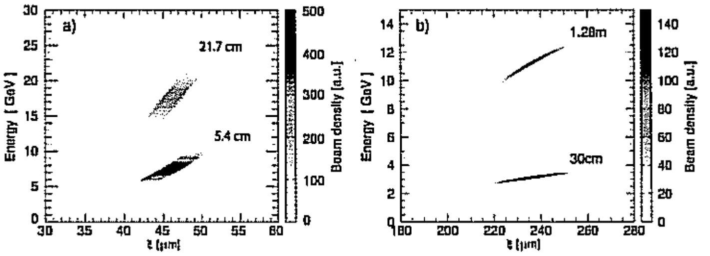

Figure 1 : QuickPIC results for (a) self-injection regime with theoretical scaling predicting 13GeV after 22 cm propagation, and (b) an ongoing simulation of the external-injection/external-guiding configuration, with theoretical scaling predicting 53GeV after 5.28m.

The same scenarios were simulated in OSIRIS in the boosted frame, since a common full PIC simulation cannot be accomplished with the computational power currently available: this would take more that 1 year on hundreds of processors. The boosted frame scheme leverages on the Lorentz transformation of the laser pulse and plasma column to reduce the numerical requirements, in particular the number of time steps involved. This enables fully kinetic three-dimensional simulation of LWFA with considerably smaller computational cost [14]. The scheme can be applied to the design of current experiments in the mm scale, allowing fast parameter scans, but can also be used to simulate new scenarios to be available with the next generation of laser systems. Fully kinetic simulations in the Lorentz boosted frame were performed for the configurations described above, and confirmed the possibility to achieve energies in excess of 10 GeV with self-injected electrons, and nearly 50 GeV for externally-injected bunches with a 300J class laser.

Designs for an LPLC [15] have indicated that accelerator stages in the 10 GeV range may be optimal for such a machine, and next generation lasers including the proposed BELLA project at LBNL will access this regime. Stages were simulated in the less strongly driven, quasi-linear regime, where dynamics are similar for electrons and positrons, using a combination of scaled, envelope, and

Lorentz boosted simulations to access the required physics. VORPAL simulations demonstrated that scaling the laser spot size and pulse length with the plasma wake wavelength allows a shorter simulation at high density using explicit PIC to deduce the properties of 10 GeV stages, and characterized the wake structure and evolution [16, 17, 18]. A wide range of laser and plasma parameters were simulated, establishing conditions for a stage that efficiently transfers laser energy into the particle bunch while maintaining high beam quality, including efficient depletion of the laser energy into the accelerating field of the wake and high transfer of laser energy deposited in the wake to the particles. These simulations showed that low energy spread 10 GeV bunches of both electrons and positrons can be obtained using a petawatt laser. In this regime, the transverse fields of the wake are also shaped by the shape of the laser spot. Simulations are now using this feature, and further tailoring of the laser and plasma, to control beam propagation and further increase achievable performance [19]. An advantage of scaled simulation using a full PIC code is that because the laser period is resolved, laser evolution up to depletion can be modeled to evaluate efficiency.

To accurately model laser and electron bunch oscillation, VORPAL envelope simulations have been used to model 0.6m of a 10 GeV stage at the design density of $10^{17}/\text{cm}^{-3}$ [10]. Fig. 2 compares the transverse profiles of the laser pulse using the envelope model and a scaled simulation; one sees that the envelope model correctly captures the self-focusing oscillations. These simulations also established designs for other applications, including Thomson gamma sources [20].

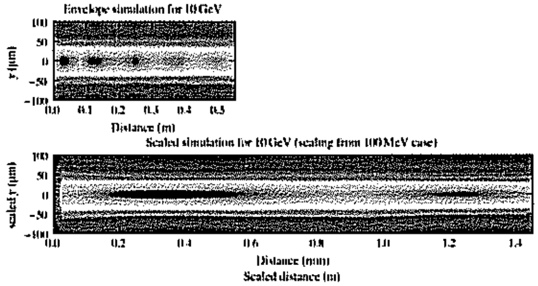

Figure 2 : Evolutions of the transverse profiles of the laser pulse for a 10 GeV LWFA stage with the envelope model (top) and the scaled simulation (bottom) using VORPAL.

Simulations of 10 GeV stages in the boosted frame are also being conducted in Warp and VORPAL. Warp simulations of scaled 10 GeV stages at plasma density $n_e=10^{19}\text{cm}^{-3}$ were performed in 2-1/2D (see [21]) and 3-D (see Fig. 3) using Lorentz boosted reference frames with relativistic factor γ between 1 (laboratory frame) and 10. The high density simulations result in short run time for effective benchmarking between the algorithms. Two figures of merit are considered (both compared in the laboratory frame): (a) the peak energy of the accelerated electron beam; (b) the average energy history of the electron beams. Agreement within a few percent is observed on the beam peak energy between calculations in all frames, with a speedup of 100 measured between the calculations in the laboratory frame (total time ~35,000 sec. ~10 hours) and in the frame at $\gamma=10$ (total time ~350 sec. ~6 min.). The average beam energy history reveals agreement at a few percents level for the accelerating phase, followed by a growing discrepancy during the decelerating phase, when using a low dispersion electromagnetic solver [14]. The agreement in the deceleration phase is much improved if the standard Yee solver is used, at a cost in CPU time, but nonetheless achieving a maximum measured speedup of 10. The gain in efficiency scales roughly as 1/n where n is the plasma density. Studies are hence in progress to use boosted frame simulations to directly simulate 10 GeV stages at plasma densities of 10^{17}cm^{-3}, which are not presently computationally accessible using conventional explicit simulations.

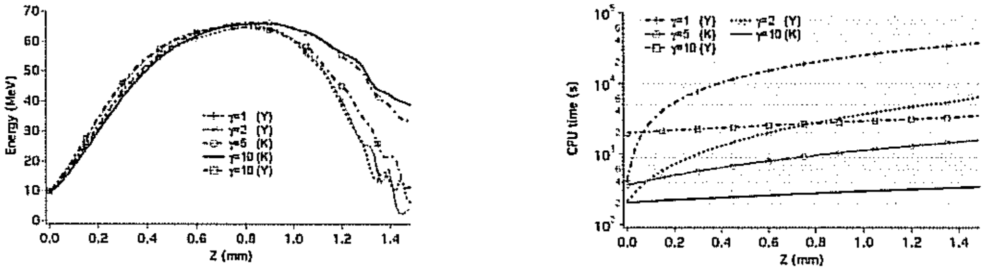

Figure 3: (left) average beam energy versus longitudinal position in the laboratory frame from 3-D Warp calculations in a frame moving at γ=1, 2, 5 and 10; (right) CPU time recorded as the beam crosses successive stations in the laboratory frame (all runs performed using 256 cores on LBNL Linux cluster). The symbols (Y) and (K) indicate respectively whether the standard Yee scheme or a low-dispersion solver was used.

For collider and many light source applications, advances are required in all of the methods to accurately model the very tight beam specifications required. Momentum errors were reduced using high order force weighting and smoothing, improving simulation of experimental momentum spread [22], and further development may be required for collider applications with extremely low emittance.

We collaborated with researchers in the VACET project and at Tech-X to produce VisIt parallel 3D renderings allowing visualization of large simulation outputs, and to improve analytics by beginning to automate detection, characterization, and fast parallel tracking of particle bunches in the simulations [18, 23, 24].

2.2. Modelling radiation emission from LWFA

The self-injected electrons in a LWFA perform betatron oscillations in the ion channel associated with the blowout region created by the laser [25, 26], and will therefore radiate. These oscillations occur as the particle is gaining energy from the wakefield, and can be observed in the trajectories depicted in Fig. 4. The radiation is in the X-ray to gamma ray range and could be used as a light source.

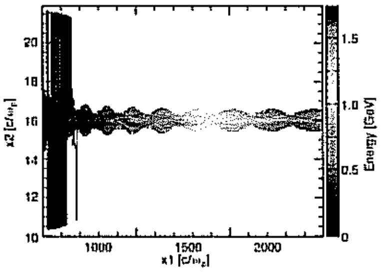

Figure 4: Projection of the trajectories (color represents energy) of the self-injected particles for the ~1.5GeV LWFA configuration. Particles (moving from left of right) are mostly injected off-axis.

By post-processing the particles trajectories, estimates for the radiation can be obtained. A detailed discussion of the radiation calculations using particle tracking in OSIRIS and the visualization package VisXD [27] will be presented in a future publication [28]. We note that, as conditions close to the energy frontier are reached, the energy losses associated with the betatron motion of the electrons in the ion channel should be considered. For these scenarios where radiation losses are important, numerical codes should include the radiation damping physics [29], as QuickPIC and OSIRIS already do. Particle tracking results in VORPAL have been used to analyze injection trajectories, formation of high quality bunches, and betatron oscillation [18, 23].

3. Recent progress and results of simulation of PWFA-LC

For the next TeV scale linear collider, high energy (500 GeV) electron/positron beams with high charge and high quality are required. One possible design based on PWFA is the PWFA-LC concept [30]. In this concept, 19 beam trains each with 250 (125×2) bunches and energy of 25 GeV are used. These beam trains are fed into 19 meter-long plasma cells and drive the plasma wakefields in each cell, respectively. The main beams that are used for the collider also have initial energies of 25 GeV. Each main beam immediately follows a drive beam in the plasma cell and extract energy from the wakefield. It is designed that the main beams will gain 25 GeV in each cell while achieving small energy spread ($< 1\%$) and emittance preservation. By staging the 19 plasma cells, the main beams will be accelerated to 500 GeV required for collision. Preliminary and non-optimized simulations show that $>50\%$ energy transfer efficiency is obtained while keeping the energy spread $< 1\%$.

The next linear collider also requires transverse beam size on the order of 10 nm to achieve high luminosity, demanding high transverse resolution in the simulations. Furthermore, 3D effects such as the hosing instability [31] and asymmetric main beam spot sizes require a 3D model. The quasi-static model is a practical choice for such simulations due to the elimination of the Courant condition. Recently the scalability of QuickPIC has been improved with a pipeline algorithm and a new 2D domain decomposition strategy for particle beams, significantly increasing the simulation throughput and enable higher transverse resolution and therefore more realistic spot sizes in the simulation.

3.1. Scaling quasi-static algorithm to 10,000+ processors

The quasi-static code QuickPIC is built on the UPIC Framework [32] with spectral space solver. A strong scaling test has shown that the UPIC Framework can scale to more than 8,000 processors for a 3D electromagnetic problem of 512×256×512 grids and 16 particles/cell on the latest homogeneous computer architecture with 2D spatial domain decomposition. The original implementation of QuickPIC with 1D spatial domain decomposition has similar but worse scaling property in comparison to the UPIC Framework. In the QuickPIC computation cycle, the plasma response to the driver is evaluated by sweeping a slice of plasma along the longitudinal direction and calculating the trajectories of the plasma particles as the slice moves from the front of the driver to the back [11, 33]. The latency in the Fourier space field solver for each update of the transverse 2D slice eventually limits the number of processors QuickPIC will scale to. For example, for $10^3 \times 10^3$ grids and 4 particles/cell in the 2D slice, the field solver can scale to 512 homogeneous processors.

In order to scale to 10,000+ processors, an algorithm that can allocate more computation tasks to the problem is needed. For a fully-explicit PIC code, this can be done through dividing the simulation domain to finer grains or in more spatial dimensions (at most 3 spatial dimensions). While for the quasi-static algorithm, a software pipelining technique can be used [34]. The pipelining algorithm adds another level of parallelism to QuickPIC by dividing the job into many pipeline stages along the direction of propagation. Each pipeline stage consists of a group of processors that works on its own domain. However, unlike conventional domain decomposition for which each processor does the calculation on its own input and communicates with its neighbors through the boundary, the initial state of the whole 2D plasma slice for a particular pipeline stage is transferred from an updated 2D slice from the previous stage (the exception is the first pipeline stage which always starts with a fresh uniform plasma slice). Therefore, pipeline stages are separated not only spatially but also in time.

The pipeline algorithm enables the efficient use of more than 1,000 processors, dramatically decreasing the turn-around time of a PWFA simulation. It has been successfully implemented into both the basic and full quasi-static versions of QuickPIC for the particle beam driver and verified to produce the same result as the non-pipelining version. The HDF I/O routine has also been modified to allow data merging of each processor group on the fly without post-processing. Performance measurement shows that the speedup of the 2D plasma solver in the pipelining mode is nearly ideal. This is because the data transfer between two successive pipeline stages is relatively inexpensive compared to the time spent on the solver itself and the transfer also overlaps with the computation. For the whole computation cycle including the 3D beam update, the overall efficiency of each processor group in the pipelining mode reaches 85% (with 2048 processors in 64 pipeline stages) relative to the non-pipelining mode. Thus using 64 pipeline stages leads to a 54 times reduction in turn-around time for a long simulation (assuming filling time of the pipeline is negligible).

The pipeline algorithm works without particular requirements on the domain decomposition, therefore one can still choose an appropriate domain decomposition strategy within each pipeline stage. In the first implementation of the pipelined QuickPIC code, 1D domain decomposition the longitudinal direction was used, same as the pipeline decomposition direction. This setup eventually limits the maximum number of processors to the number of grid points available in the longitudinal direction which is typically on the order of 1000. Recently, we further improved the pipeline algorithm with a 2D domain decomposition in the pipeline stages. Preliminary scaling tests show that it will scale to at least 16,384 processors for a PWFA simulation with 2048×2048×256 grid points (Figure 5).

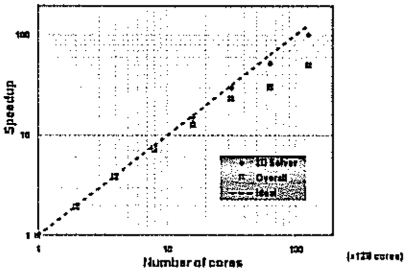

Figure 5 : Strong scaling of QuickPIC on the NERSC Franklin platform with 2048×2048×256 grids and 4 particles/cell. 128 cores per pipeline stage are used and 1 to 128 stages are tested. The smallest domain size is 2048×16×2. The blue diamonds and the red squares are measured speedup for the calculation of a plasma slice in the 2D solver and for one 3D step, respectively.

3.2. High resolution simulation of the PWFA-LC concepts

A PWFA stage that can provide a large acceleration gradient and high total efficiency (30% ~ 90% from drive beam to main beam) and accelerate sufficient charge (~10^{10} electrons) with high beam quality and low energy spread (0.1% ~1%) is essential to the PWFA-LC concept. An ideal design for the first stage (25GeV) and the last stage (475GeV) of the PWFA-LC was investigated using QuickPIC [35]. In this design, the beam loading theory [36] based on the theoretical framework [37] of nonlinear wake excitation was used to derive the optimal current profiles of the drive and main beams. In the simulation, the main beam with 1.73×10^{10} electrons gains 25 GeV in 0.6 meter with an acceleration gradient of 42.7 GeV/m. Less than 1% energy spread of the main beam was achieved and the efficiency of energy transfer from the drive beam to the main beam is 51%.

Using the QuickPIC code with the latest improvements in simulation capability described in section 3.1, a more detailed study of a PWFA-LC stage with main beam parameters close to the requirements of a future TeV collider can be carried out. For such a matched beam with tight spot sizes, the peak beam density would be 10^3~10^4 times higher than the plasma density, possibly causing the plasma ions to move even on the sub-picosecond time scale of the beam duration [38]. This issue also needs to be considered for LWFA collider design not using a weakly nonlinear wake. The ions are attracted to the center of the beam, therefore, to include the ion dynamics in the simulation also demands a transverse resolution of ~10 nm. As a first step, a preliminary high resolution simulation of a nominal PWFA-LC stage has been conducted. The drive and main beams have 2.9×10^{10} electrons and 1×10^{10} electrons respectively and they are both modeled with ~4,200,000 macro-particles. Their initial energies are 25 GeV. The emittance of the main beam is 0.093 mm·mrad and the matched spot sizes are 100 nm for $n_e=1\times10^{17}$ cm^{-3}. The drive beam spot size can be larger as it is not used in the final collision. An emittance of 10 mm·mrad is chosen which is typical for current state-of-the-art electron beams from a linac. The matched spot size is 1 μm. The resolution of the simulation is 49nmx49nmx304nm and it is carried out with 8192x8192x1024 grid points using 8192 processors on Franklin. There are 4 particles per cell for the plasma electron and ion respectively. Their separation is comparable to the real atom separation of ~20 nm at 1×10^{17} cm^{-3}. We note that this is the first time in PIC simulation of PWFA that one can simulate nearly all the particles in a real plasma. Figure 6 shows the results from the simulation.

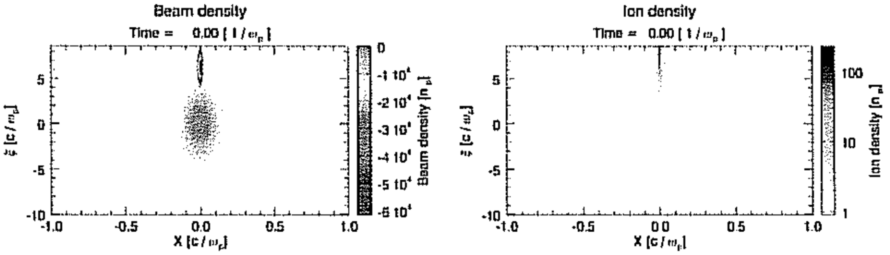

Figure 6 : High resolution QuickPIC simulation of the PWFA-LC concept with mobile ions. The spot sizes of the beams are matched to the plasma density. Left: the drive and the main beam (moving down); right: the density of the ion background showing ion concentration near the axis.

4. Future challenges of PIC simulations and the path towards the energy frontier

The recent progress on the development of the boosted frame technique and reduced models (envelope model for LWFA only and quasi-static model for LWFA and PWFA) and the usage of massive parallel computing resources have enabled simulation studies of the near term LWFA/PWFA experiments and conceptual designs of the next generation facilities. However, the needs to faithfully simulate the laser/beam driver evolution and the main beam dynamics in a future plasma-based collider are still challenging. The computation cost of the former simulation can be reduced by using the boosted frame technique or envelope/quasi-static model. Here we estimate the computation requirements for the main beam. For example, the smallest emittance of the main beam in a TeV collider is 0.04 mm·mrad. The matched spot size of the main beam at 500 GeV in a plasma of 1×10^{17} cm^{-3} will be 30 nm, which is three orders of magnitudes smaller than the longitudinal spot size or the plasma wavelength. Furthermore, at the final focus where a plasma lens may be used to focus the beam to the interaction point, the smallest beam size in the transverse dimension is 6 nm requiring nm resolution in the simulation. However, the transverse box size needs to be around 20 collisonless skin depths, which is ~300 microns. This is 10^5 times larger than the required resolution. In the longitudinal direction, the plasma wavelength needs to be well resolved using O(1000) grids. Therefore a realistic 3D high-resolution simulation of the accelerated beam would need $10^5\times10^5\times1000 = 1\times10^{13}$ grids and 4×10^{13} particles (assuming 4 particles/cell), demanding extremely large amount of memory and processors. The time step and the number of time steps for meter-long propagation distance are ~ 0.002 fs and 1×10^9 for a full PIC simulation, or ~ 16 ps and 208 for a quasi-static simulation. Such simulation is difficult to run even using the quasi-static model and the largest computers in the world. Alternatively, as described above, in the linear regime, the focusing fields can be shaped, and simulations are investigating how this can allow use of larger diameters beams [19].

To increase the capability of large-scale PIC simulation for the energy frontier, a comprehensive approach needs to be taken. This will include algorithm development to enhance the single core performance of the PIC algorithm and to optimize PIC codes to scale efficiently on many-core system and to adapt them to heterogeneous architecture. On the other hand, mesh refinement, further development of the boost frame algorithm and the envelope/quasi-static models might also play important roles to reduce the computational demands substantially.

The suit of PIC codes in the SciDAC COMPASS project are constantly being improved to run efficiently on latest homogenous platforms. For example, OSIRIS performs well on a single processor while maintaining high parallel efficiency with load-balancing, i.e., >85% for strong scaling and >95% for weak scaling on greater than 32,000 processors of the Argonne BlueGene Intrepid Computer (Figure 7). Additional work on performance also focuses on better utilization of the vector unit of modern CPUs. Preliminary work has shown a 2-2.5x speedup on the Nehalem architecture with vectorized field interpolation, particle push and current deposit routines.

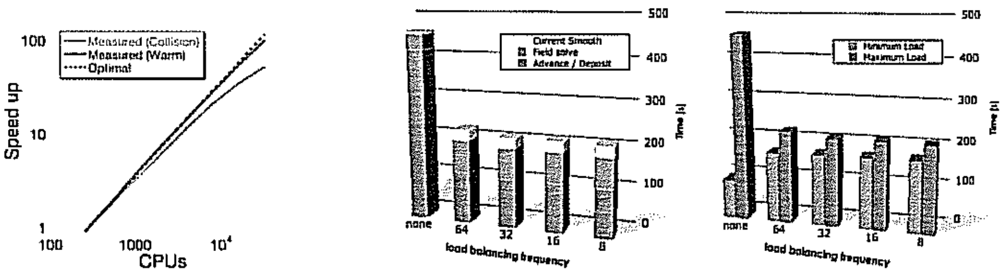

Figure 7: Strong scaling for OSIRIS on up to 32,768 processors on the Intrepid BlueGene machine (left) and timing results (middle and right) vs. load balance frequency when the particle push and field solve are load balanced together.

However, near-term petascale platforms will likely make a paradigm shift to the heterogeneous architecture with millions threads and processing units. One has to ensure that the key PIC algorithms will continue to scale well on such a platform. Currently, efforts are being made at UCLA to rewrite critical pieces of the PIC algorithm to take advantage of new architectures such as GPUs. As a first step, the kernel of the UPIC framework has been ported to the GPU and optimization of the key routines is under way. On the other hand, the latest version of OSIRIS already supports arbitrary precision floating points (which allows the data to be moved to and from GPU's).

In addition, the Adaptive Mesh Refinement (AMR) technique can be used to significantly reduce the requirements for computer memory and the number of operations for large scale simulations. Its application to the fully self-consistent modeling of beams and plasmas is especially challenging, due to properties of the Vlasov-Maxwell system of equations, but successful PIC simulations using AMR have been reported [39]. The application of AMR to the modeling of LWFA is being explored with Warp. For the simulation of a 10GeV LWFA stage, the wake wavelength is O(100μm) while the electron bunch and laser wavelength are typically submicron in size. As a result, the resolution required for different parts of the problem may vary by more than two orders of magnitude in each direction, corresponding to up to 6 orders of magnitude of possible (theoretical) savings by use of mesh refinement. While algorithm limitations will most probably limit the actual speedup to lower values, we anticipate that speedups of one order of magnitude or more are achievable. We note that the savings offered by mesh refinement will apply in addition to the savings provided by the various techniques already employed (envelope, quasi-static, fluid, boosted frame).

5. Conclusions

The plasma-based advanced accelerator is a promising technology to reduce the size and cost of a linear collider or a X-ray light source by providing orders of magnitude higher accelerating gradient than the conventional RF technology. Large scale PIC simulations have provided insight into the proof-of-principle LWFA/PWFA experiments and elucidated the relevant physics. Recently, developments of various PIC models and techniques, such as the implementation of the envelope model and boosted frame approach, the enhancement of the quasi-static model with a pipelining algorithm and the performance and scaling optimization of the full PIC model, pave the way of using large-scale PIC simulations to model the next generation LWFA/PWFA experiments with PW class lasers or state-of-the-art electron beam drivers. Conceptual designs of 10 GeV LWFA module or 25 GeV PWFA module are now being developed and refined with PIC simulations as a guide. In the future, adaptation of PIC codes to the latest vector units in a modern processor and innovative many-core/heterogeneous computing architecture and algorithm development such as the AMR technique will be pursued to meet the simulation requirement for the energy frontier of 1 TeV or more.

6. Acknowledgements

This work is supported by Department of Energy, Office of Science, through the SciDAC program, and by NNSA through NA-22. Simulations are conducted at the NERSC platforms, the UCLA Dawson cluster, the LBNL Linux cluster, the LLNL ATLAS facility and the Argonne ALCF facility.

References

[1] T. Tajima, J. M. Dawson, Phys. Rev. Lett., 43, 267 (1979).
[2] P. Chen et al, Phys. Rev. Lett., 55, 1537 (1985).
[3] C. K. Birdsall et al., Plasma Physics via Computer Simulation, McGraw-Hill, New York, 1985.
[4] J. M. Dawson, Rev. Mod. Phys., 55, 403 (1983).
[5] Nieter et al., Journal of Computational Physics, 2004.
[6] R. A. Fonseca et al., Lec. Notes in Comp. Sci., 2329, 342 (2002); R. G. Hemker, Ph.D. Thesis, UCLA (2000).
[7] D. P. Grote, A. Friedman, J.-L. Vay, I. Haber, AIP Conf. Proc. 749, 55 (2005).
[8] K. Paul et al., AIP Conf. Proc. 1086, 315 (2009).
[9] J.-L. Vay, Phys. Rev. Lett. 98, 130405 (2007).
[10] B. Cowan et al., AIP Conf. Proc. 1086, 208 (2009).
[11] C. Huang et al., J. Computational Phys. 217, 658 (2006).
[12] S. F. Martins et al., in preparation.
[13] W. Lu et al, Phys. Rev. STAB, 10, 061301 (2007).
[14] J.-L. Vay et al., These proceedings.
[15] C.B. Schroeder et al., AIP Conf. Proc. 1086, 309 (2009).
[16] E. Cormier-Michel et al., AIP Conf. Proc. 1086, 297 (2009).
[17] C.G.R. Geddes et al., Proc. Particle Accel. Conf., Vancouver, Canada (2009) WE6RFP075.
[18] C.G.R. Geddes et al, SciDAC review, accepted.
[19] E. Cormier-Michel et al., in preparation.
[20] C.G.R. Geddes et al., Proc. Conf. on Applications of Accelerators in Research & Industry 2008.
[21] J.-L. Vay et al., Proc. Particle Accel. Conf., Vancouver, Canada (2009) TU1PBI04.
[22] C. G. R. Geddes et al., AIP Conf. Proc. 1086, 12 (2009).
[23] O. Ruebel et al., Proc. Supercomputing Conference 2008.
[24] D.M. Ushizima et al., International Conference on Machine Learning & Applications, 2008.
[25] D. H. Whittum et al., Phys. Rev. Lett. 64, 2511 (1990).
[26] S. Kneip et al., Phys. Rev. Lett. 100, 105006 (2008).
[27] R A Fonseca et al., Plasma Physics and Controlled Fusion 50, 124034 (2008).
[28] J. L. Martins et al., in preparation.
[29] S. Kiselev et al., Phys. Rev.Lett., 93, 135004 (2004).
[30] A. Seryi et al., Proc. Particle Accel. Conf., Vancouver, Canada (2009), WE6PFP081.
[31] C. Huang et al., Phys. Rev. Lett. 99, 255001 (2007).
[32] V. K. Decyk, Computer Phys. Comm. 177, 95 (2007).
[33] C. Huang et al., J. Phys.: Conf. Ser. 46 190-199, 2006.
[34] B. Feng et al., Journal of Computational Physics, In Press, DOI: 10.1016/j.jcp.2009.04.019.
[35] C. Huang et al., Proc. Particle Accel. Conf., Vancouver, Canada (2009), WE6RFP097.
[36] M. Tzoufras et al., Phys. Rev. Lett. 101, 145002 (2008).
[37] W. Lu et al., Phys. Rev. Lett. 96, 165002 (2006).
[38] J. B. Rosenzweig et al., Phys. Rev. Lett. 95, 195002 (2005).
[39] J.-L. Vay et al., Phys. Plasmas 11 (2004).