Software pipeline in QuickPIC

Chengkun Huang
UCLA

Collaborators

V. K. Decyk, W. B. Mori

B. Feng, T. Katsouleas

J. H. Cooley, T. M. Antonsen

J. Vieira, L. O. Silva

Challenge in PIC modeling

Typical 3D high fidelity PWFA/LWFA simulation requirement

	Feature	Grid size limit	Time step limit	Total time of simulation per GeV stage (node-hour)*
PWFA	Full EM PIC	~0.05c/ω_p	$\Delta t < \frac{1}{\sqrt{3}} 0.05\omega_p^{-1}$	1500
	Quasi-static PIC	~0.05c/ω_p	$\Delta t < 0.05\omega_\beta^{-1} = 0.05 \times \sqrt{2\gamma}\omega_p^{-1}$	4 (20)

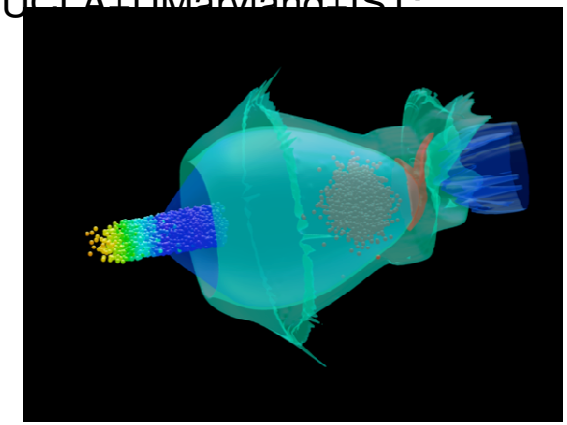
	Feature	Grid size limit	Time step limit	Total time of simulation per GeV stage (node-hour)
LWFA	Full EM PIC	~0.05 λ	$\Delta t < 0.05\ \omega_0^{-1}$	~ 500000
	Ponderomotive Guiding center PIC	~0.05c/ω_p	$\Delta t < 0.05\omega_p^{-1}$	~1500
	Quasi-static PIC	~0.05c/ω_p	$\Delta t < 0.05\ \tau_r$	~ 10

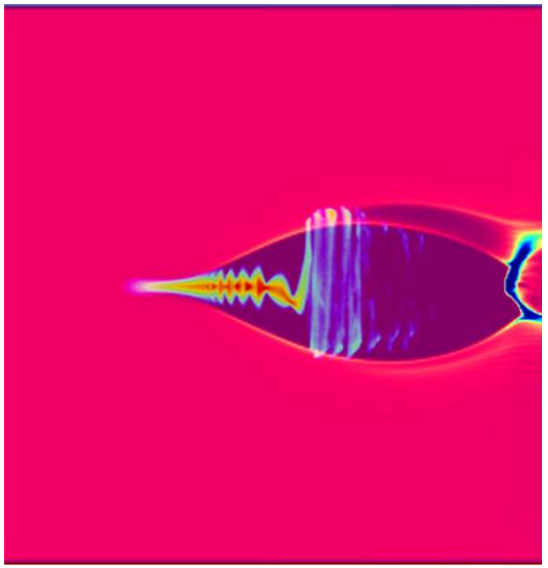
*These are rough estimates and represent potential speed up. In some cases we have not reached the full potential. In some cases the timing can be reduced by lowering the number of particles per cell etc.

QuickPIC

QuickPIC

- Massivelly Parallel, 3D Quasi-static particle-in-cell code
- Ponderomotive guiding center for laser driver
- 100-1000+ savings with high fidelity
- Field ionization and radiation reaction included
- Simplified version used for e-cloud modeling
- Developed by the UCLA+UMaryland+IST;





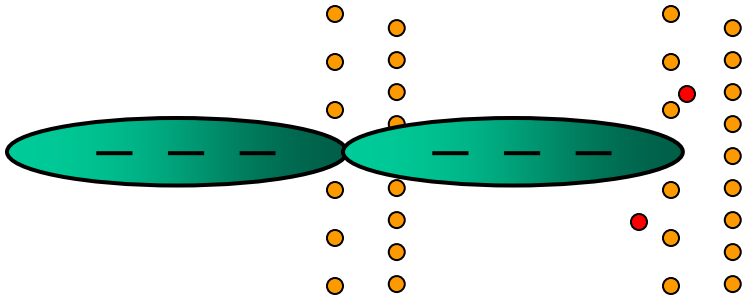
New Features

- Pipelining
- Parallel scaling to 1000+ processors

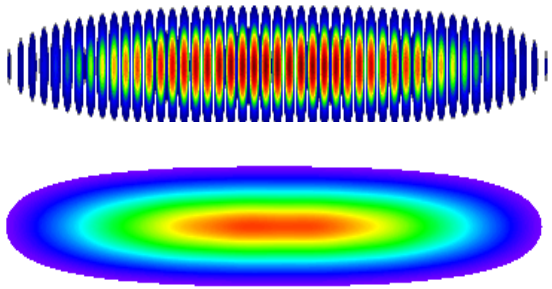
Chengkun Huang:
huangck@ee.ucla.edu
http://exodus.physics.ucla.edu/
http://cfp.ist.utl.pt/golp/epp

Quasi-static Model

- There are two intrinsic time scales, one fast time scale associated with the plasma motion and one slow time scale associated with the betatron motion of an ultra-relativistic electron beam.
- Quasi-static approximation eliminates the need to follow fast plasma motion for the whole simulation.
- Ponderomotive Guiding Center approximation: High frequency laser oscillation can be averaged out, laser pulse will be represented by its envelope.

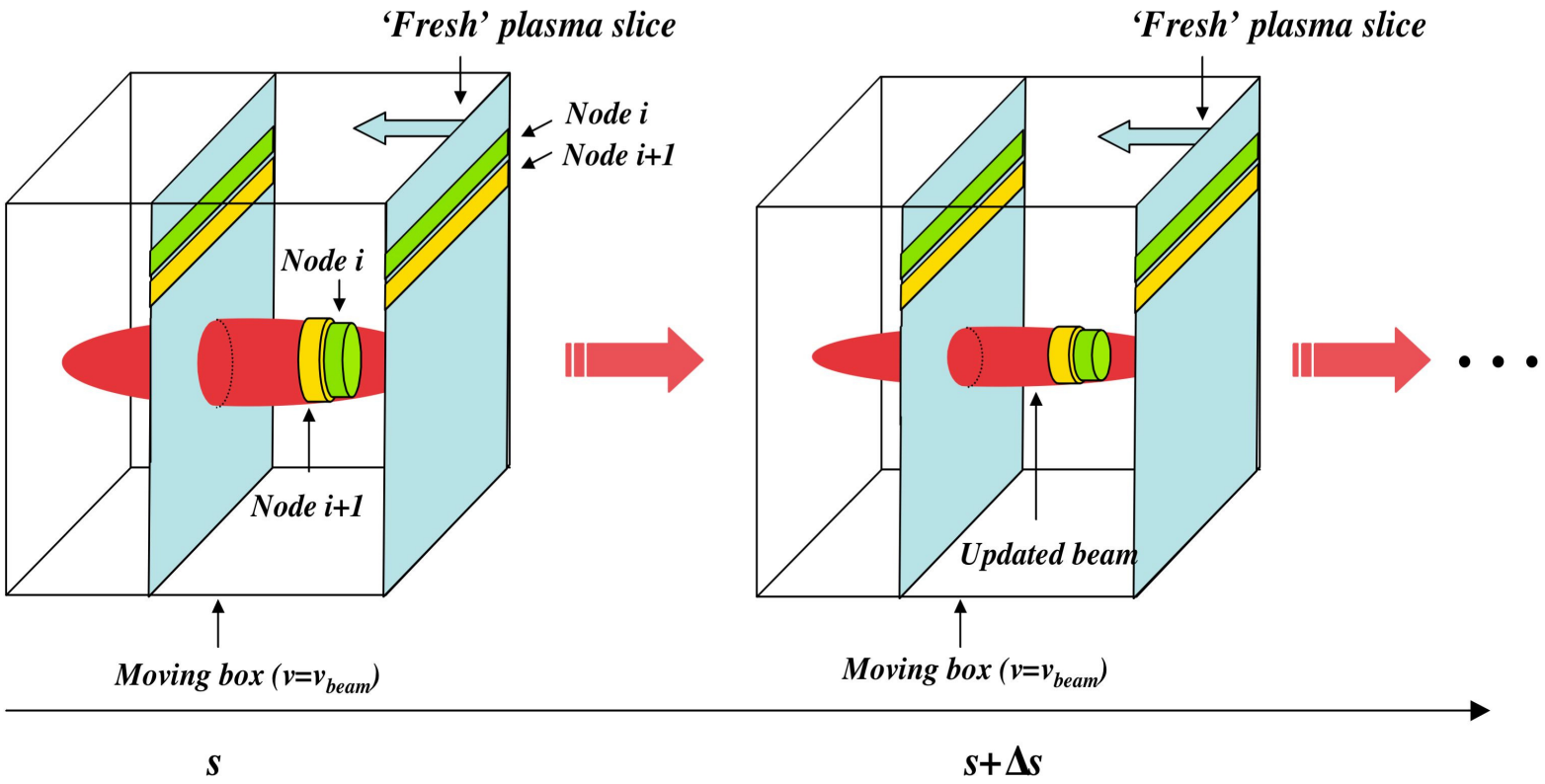


Quasi-static



Ponderomotive Guiding Center

Implementation

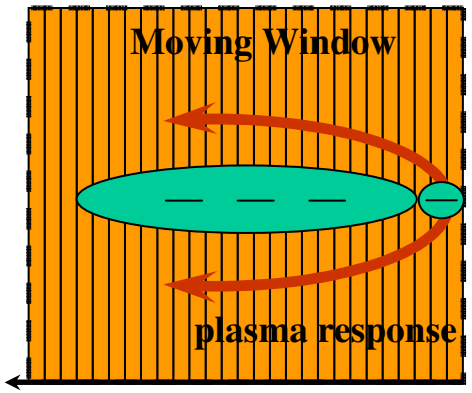


The driver evolution can be calculated in a 3D moving box, while the plasma response can be solved for slice by slice with the longitudinal index being a time-like variable.

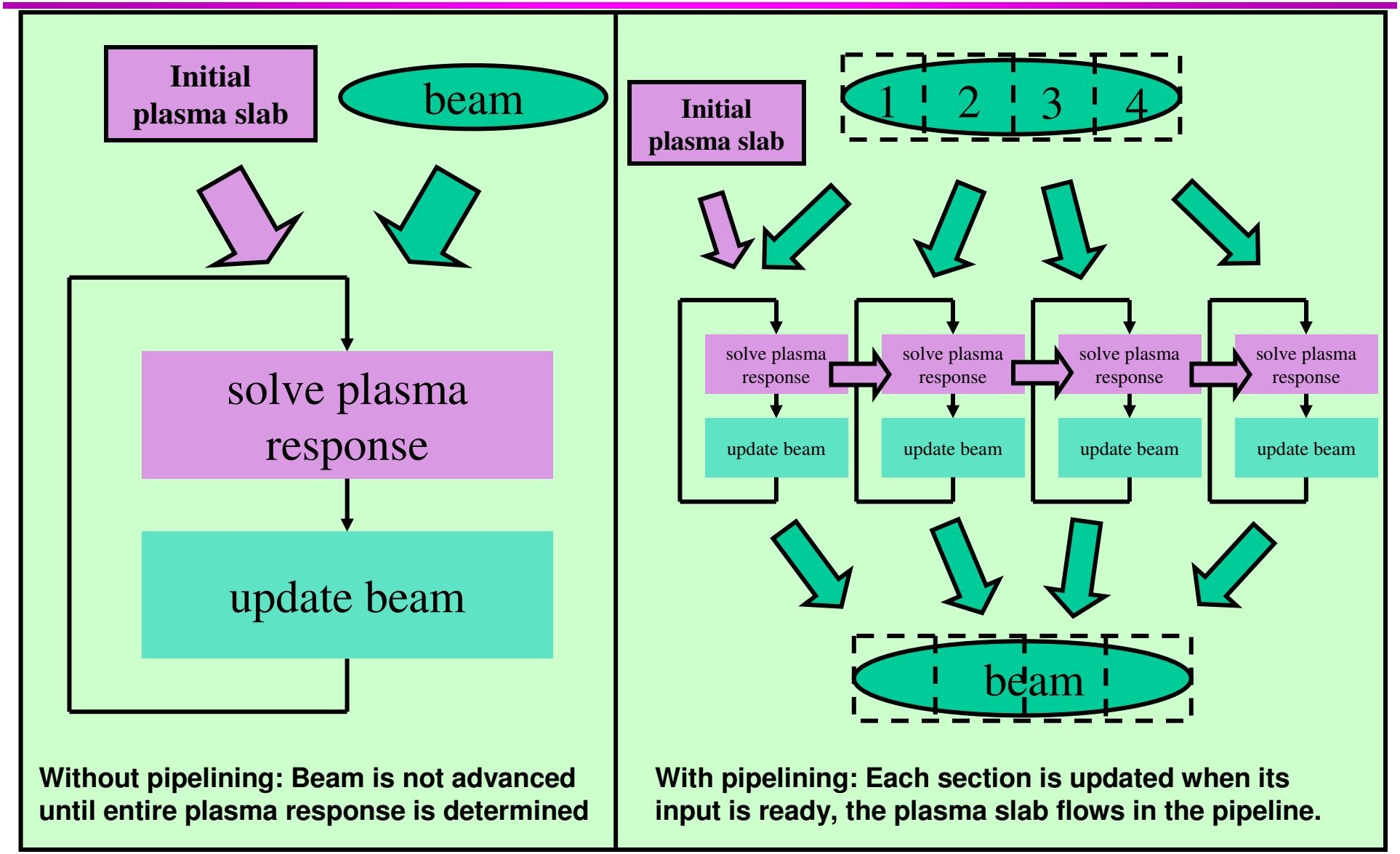
Exploiting more parallelism: Pipelining

- Pipelining technique exploits parallelism in a sequential operation stream and can be adopted in various levels.
- Modern CPU designs include instruction level pipeline to improve performance by increasing the throughput.
- In scientific computation, software level pipeline is less common due to hidden parallelism in the algorithm.
- We have implemented a software level pipeline in QuickPIC.

	Instruction pipeline	Software pipeline
Operand	Instruction stream	Plasma slice
Operation	IF, ID, EX, MEM, WB	Plasma/beam update
Stages	5 ~ 31	1 ~(# of slices)



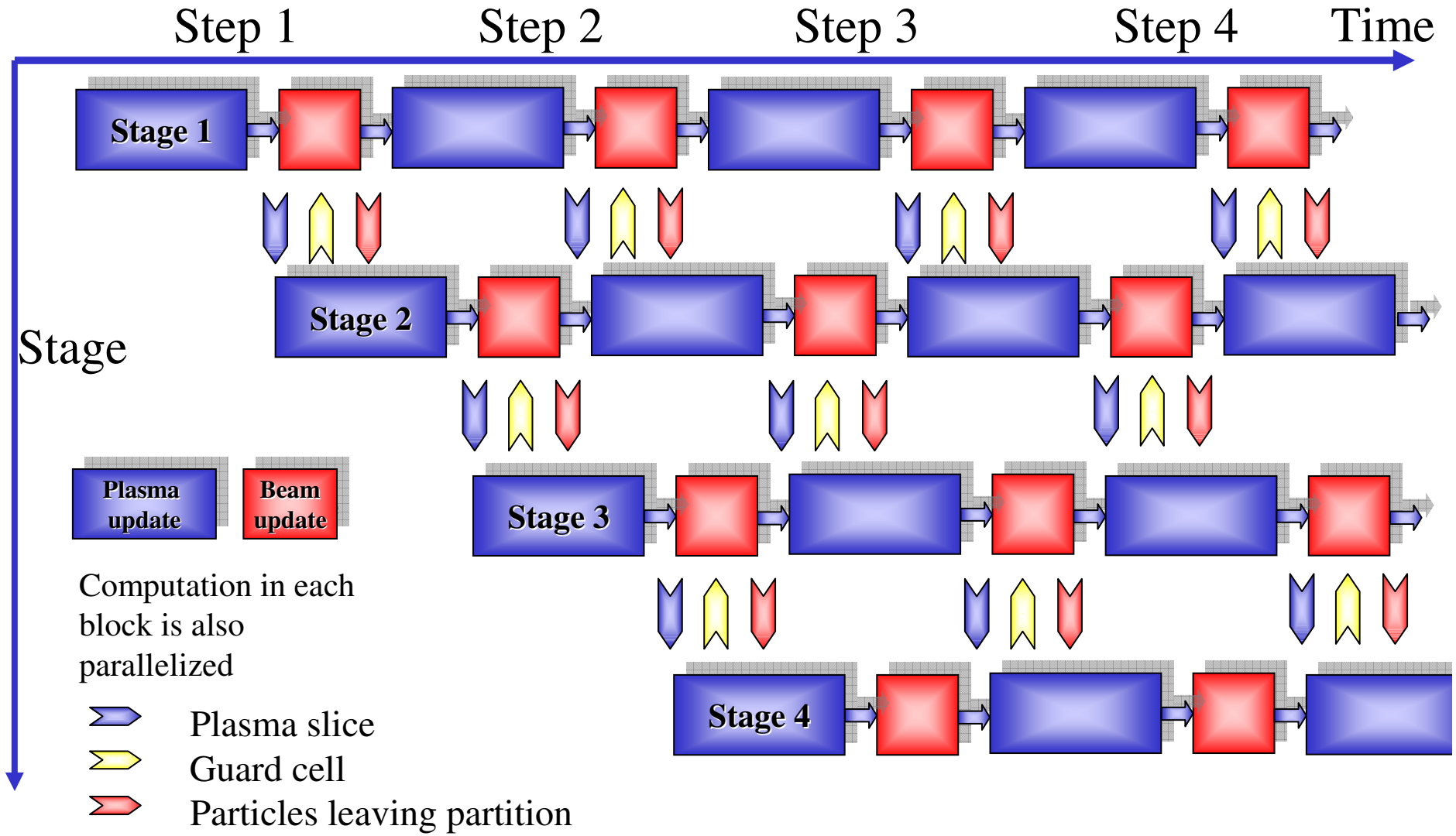
Pipelining: scaling QuickPIC to 10,000+ processors



Without pipelining: Beam is not advanced until entire plasma response is determined

With pipelining: Each section is updated when its input is ready, the plasma slab flows in the pipeline.

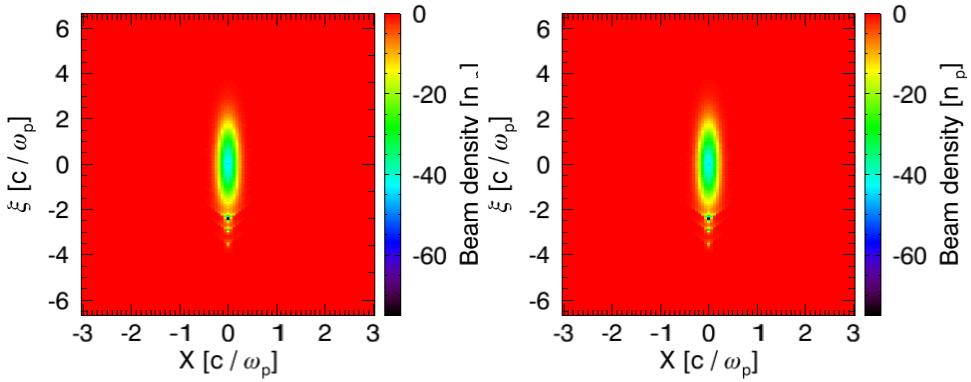
More details



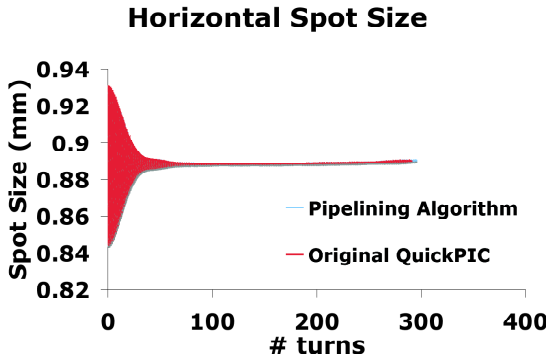
Code Verification

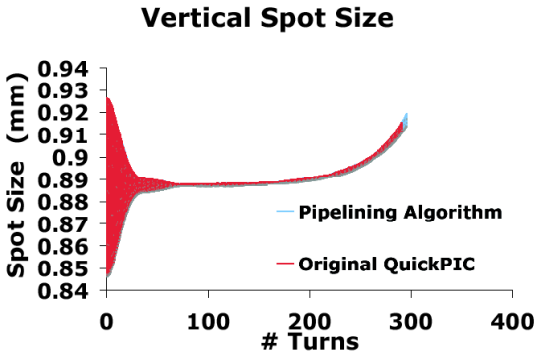
original QuickPIC

with pipelining algorithm



Electron beam densities in the PWFA simulations at the 150th time step





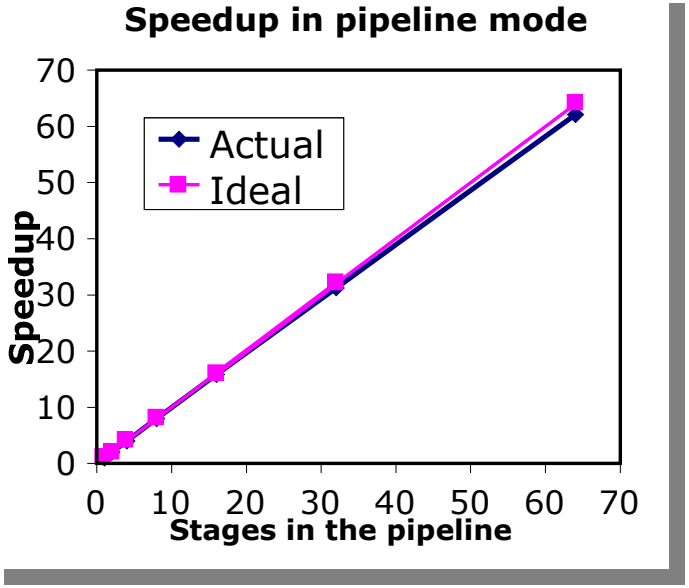
Horizontal and vertical spot sizes of the beam in basic quasistatic QuickPIC ecloud simulation

Feng et al, submitted to JCP

Scaling to 1,000+ processors

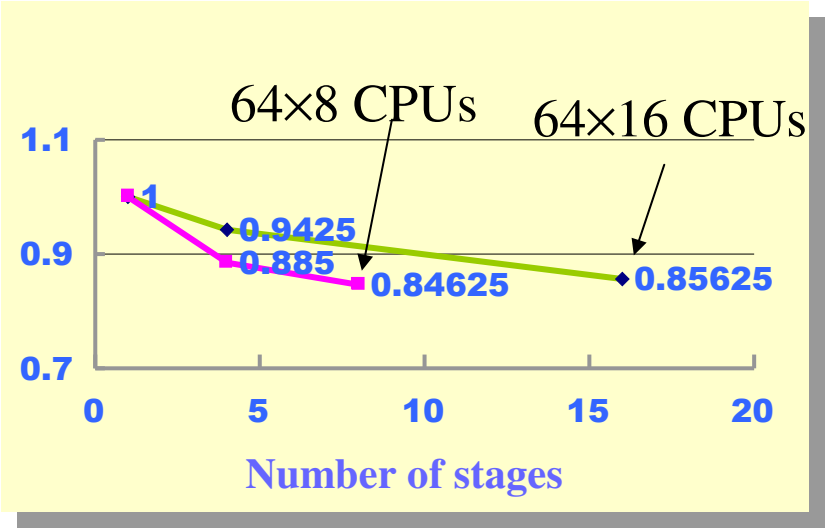
2D pipeline performance

- Near ideal speed up for 2D module in pipeline mode
- Data transfer between successive stages is inexpensive and independent of the number of stages.
- Data transfer overlaps with computation.
- Speedup drops when overhead becomes comparable to computation time.

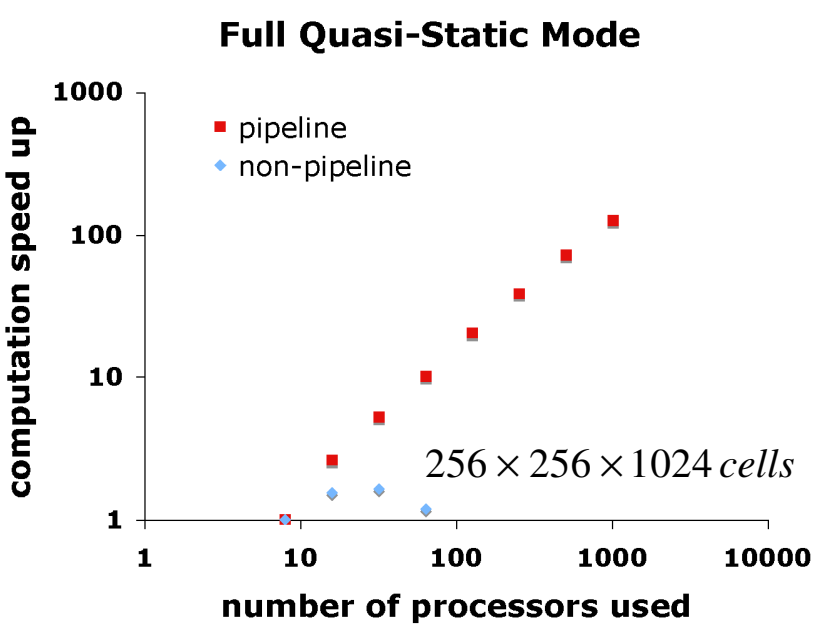


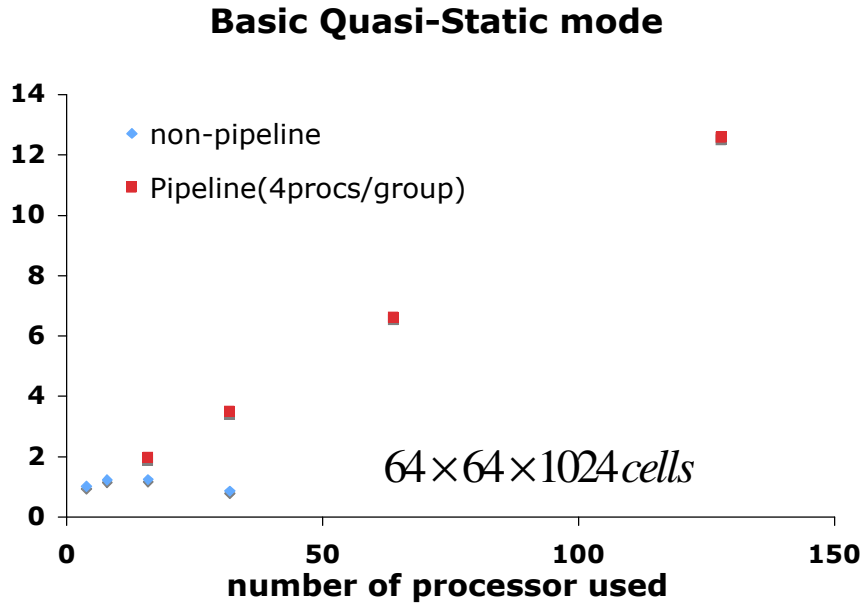
Overall performance

- 16 stages pipeline is 85% efficient compared with single (no) pipeline calculation
- 13.6 times throughput improvement with 16 stages pipeline.
- Relatively independent of # of pipeline



Performance in pipeline mode





Feng et al, submitted to JCP

- Fixed problem size, strong scaling study, increase number of processors by increasing pipeline stages
- In each stage, the number of processors is chosen according to the transverse size of the problem.
- Benchmark shows that pipeline operation can be scaled to at least 1,000+ processors with substantial throughput improvement.

Enhancing pipeline operation

Work in progress & Future improvement

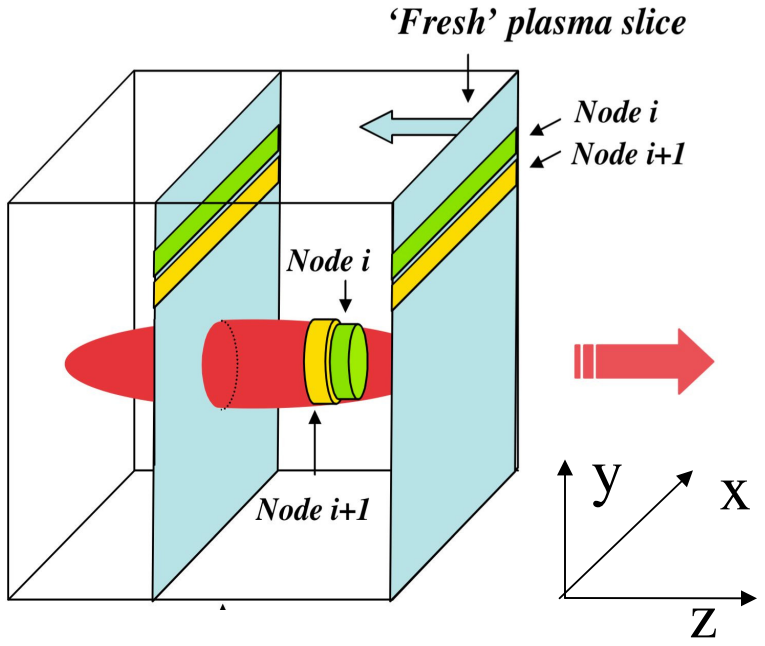
- Enhanced domain decomposition (in Z -> in Y, or 2D decomposition) for better job allocation. Also enables us to use more processors in the transverse direction and extremely small cell size which is required for simulating narrow trailing beam in PWFA-LC.

- For small pipeline stage, 3D beam update should be load balanced.

  For example, 16 pipeline stages,

  2D: 1102.8 sec / 79.4 sec = 13.9 times faster

  3D: 15.5 sec / 2.2 sec = 7.04

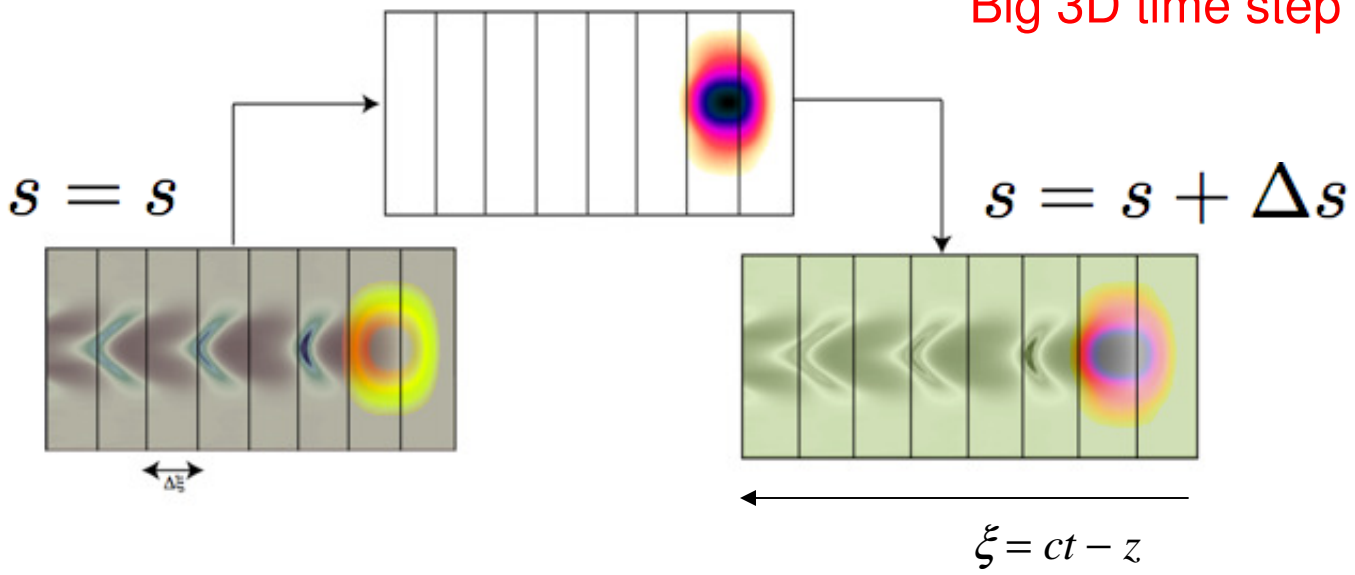


Pipelining for laser solver

$$\left[2\partial_s\left(-ik_o-\partial_\xi\right)-\nabla_\perp^2\right]\hat{\mathbf{A}}^{\mathbf{laser}}=\frac{4\pi}{c}\mathbf{J}$$

Ponderomotive guiding center approximation:

Big 3D time step



Plasma evolution:

Maxwell's equations Lorentz Gauge

Quasi-Static Approximation

$$-\nabla_\perp^2\phi = 4\pi\rho$$

$$-\nabla_\perp^2\mathbf{A} = 4\pi\mathbf{J}$$

Summary

- **QuickPIC enables scientific discovery in plasma-based acceleration by allowing 100-1000 times time-saving for simulations of state-of-art experiments and for exploring parameter space which are not easily accessible through conventional PIC code.**

- **We are working to scale QuickPIC to the petascale platform using the software pipelining technique. Initial benchmark shows very promising performance enhancement.**