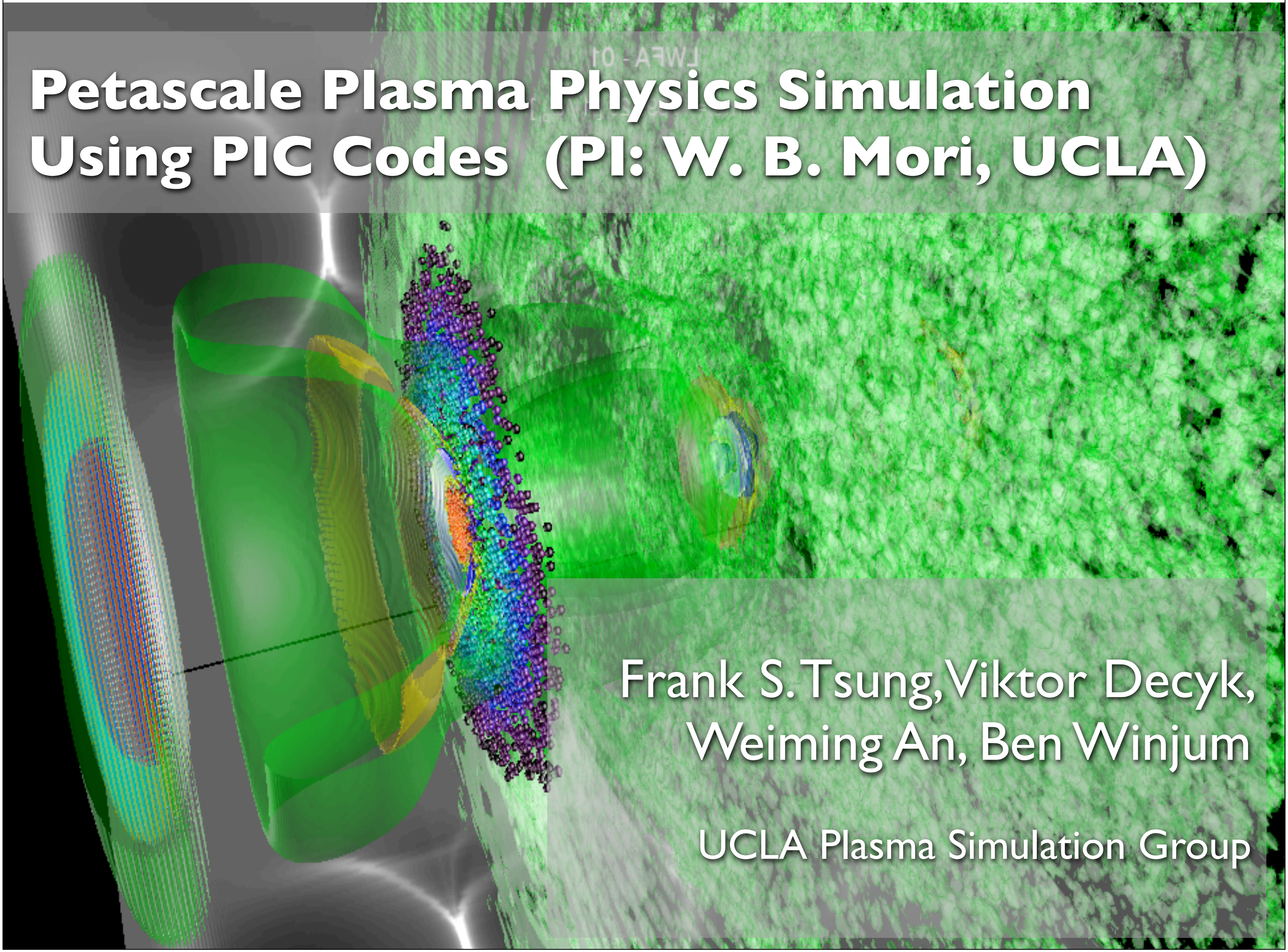
Petascale Plasma Physics Simulation Using PIC Codes (PI: W. B. Mori, UCLA)



Frank S. Tsung, Viktor Decyk, Weiming An, Ben Winjum

UCLA Plasma Simulation Group

Summary and Outline

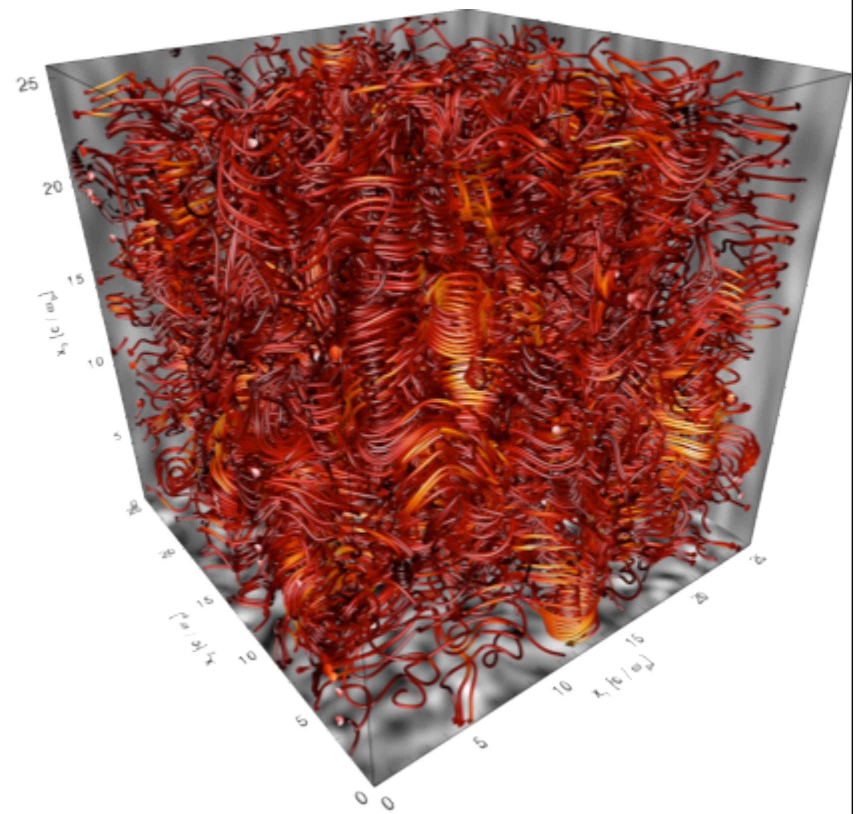
OUTLINE

- Overview of the project
  - Particle-in-cell codes
  - OSIRIS and recent >2PFlops benchmark on Blue Waters

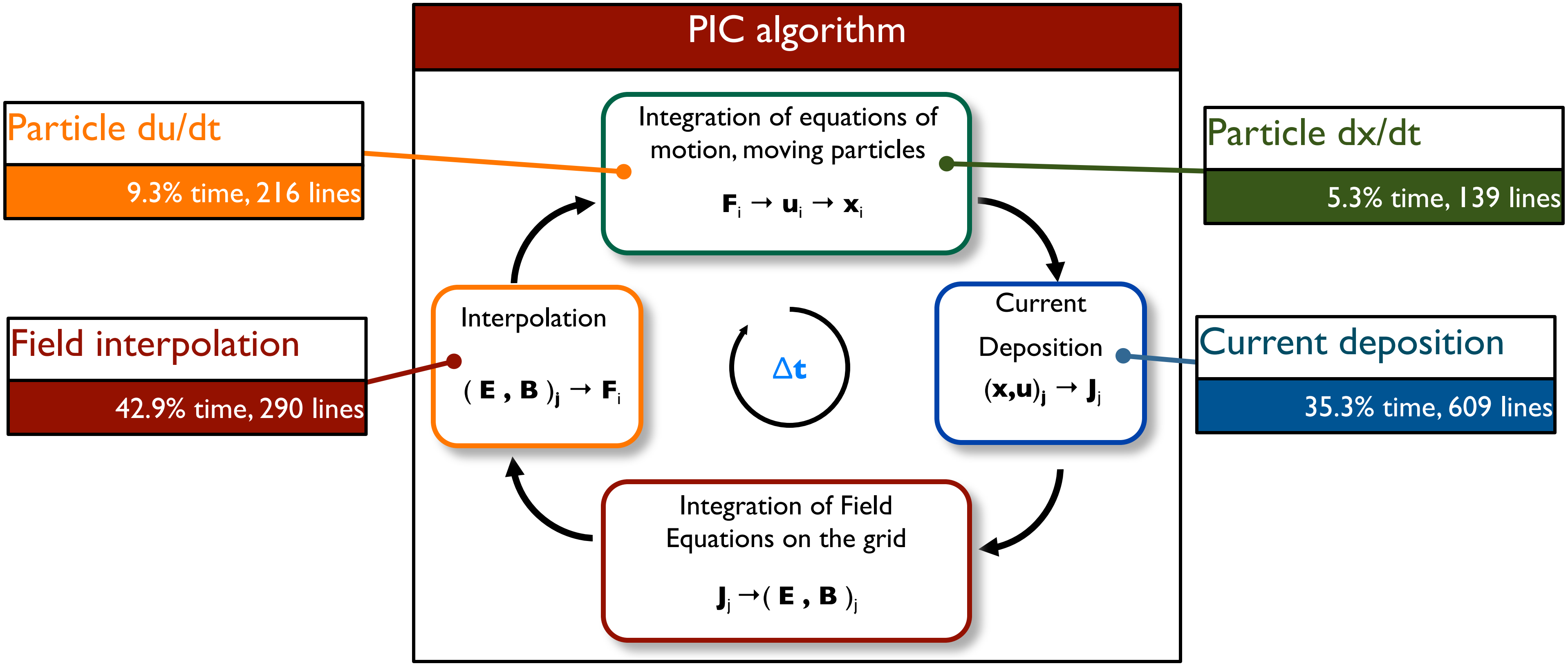
- Application of OSIRIS to plasma based accelerators:
  1. OSIRIS simulations of ring formation in Callisto experiments @ LLNL
  2. QuickPIC simulations of Ion motions using ILC beam parameters.

- Applications of OSIRIS to LPI's Relevant to IFE
  - SRS in indirect drive IFE targets (such as NIF).
  - Estimates of large scale LPI simulations (& the need for exascale supercomputers)

- Development works for Blue Waters and beyond (including GPU's and other emerging architectures) + the PICKSC Center @ UCLA



Profile of OSIRIS/Introduction to PIC



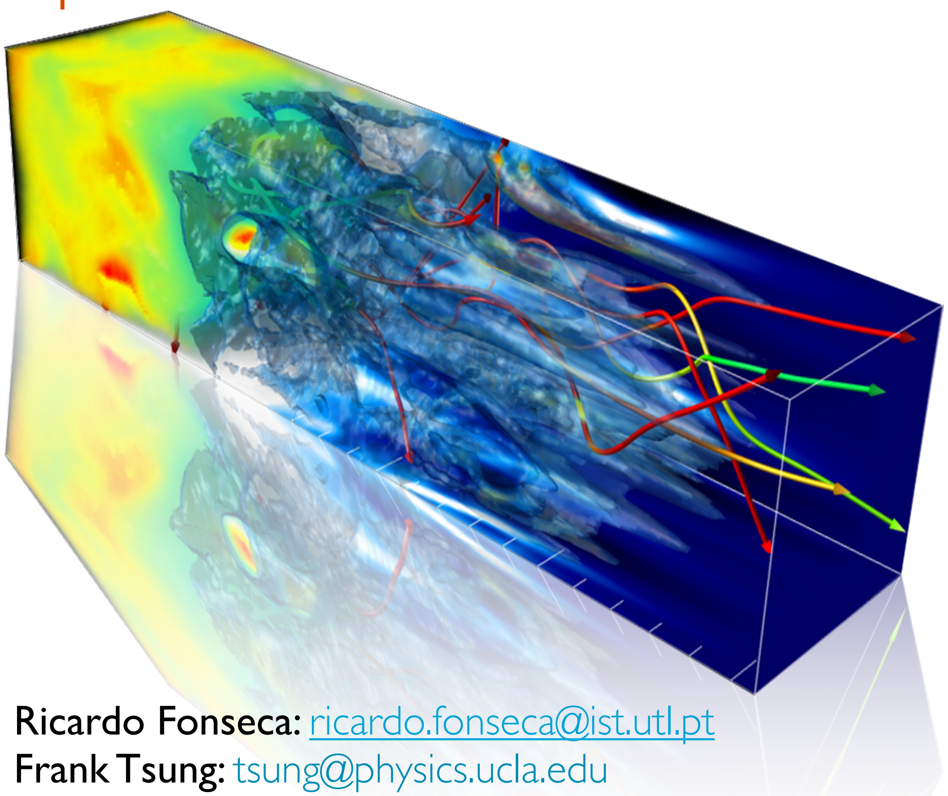
- The particle-in-cell method treats plasma as a collection of computer particles. The interactions does not scale as N^2 due to the fact the particle quantities are deposited on a grid and the fields are solved on these grids only. Because (# of particles) >> (# of grids), the timing is dominated by the orbit calculations
- The code spends over 90 % of execution time in only 4 routines
- These routines correspond to less than 2 % of the code, so optimization is fairly straight forward

osiris

osiris framework

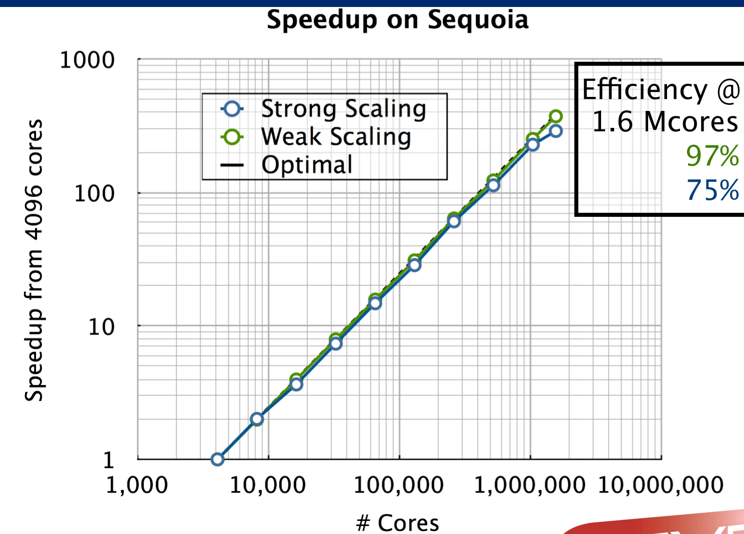
- Massivelly Parallel, Fully Relativistic Particle-in-Cell (PIC) Code
- Visualization and Data Analysis Infrastructure
- Developed by the osiris.consortium

⇒ UCLA + IST



Ricardo Fonseca: ricardo.fonseca@ist.utl.pt
Frank Tsung: tsung@physics.ucla.edu

http://cfp.ist.utl.pt/golp/epp/
http://exodus.physics.ucla.edu/

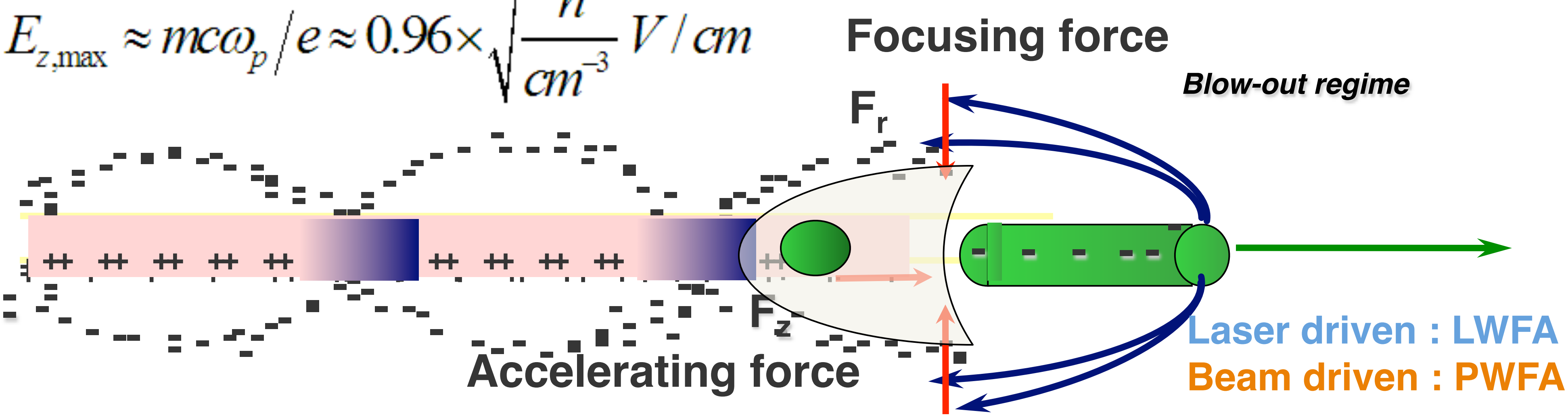


New Features

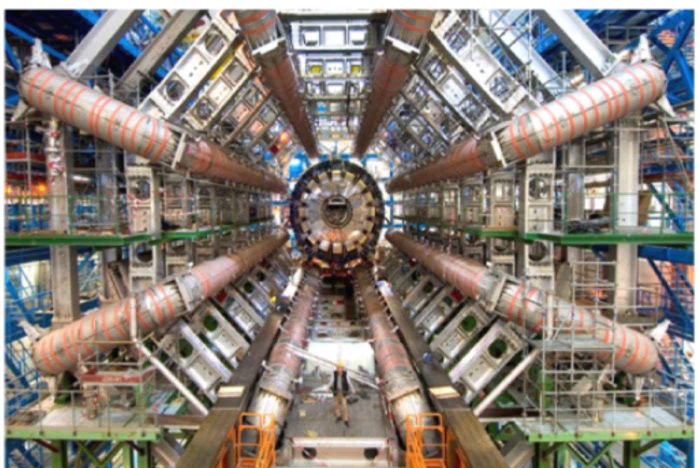
- Bessel Beams
- Binary Collision Module (to study plasmas which behave more like fluids)
- **Energy Conserving Algorithm**
- **Multi-dimensional Dynamic Load Balancing**
- **OpenMP/MPI hybrid parallelism**
- PML absorbing BC
- **Higher order splines**
- Parallel I/O (HDF5)
- Gridless cylindrical mode
- > 2.2 PFlops on Blue Waters & good scaling on > 1.5 million cores (Sequoia supercomputer @ LLNL)

Plasma-based accelerators and applications

$$E_{z,\max} \approx mc\omega_p / e \approx 0.96 \times \sqrt{\frac{n}{cm^{-3}}}\, V/cm$$

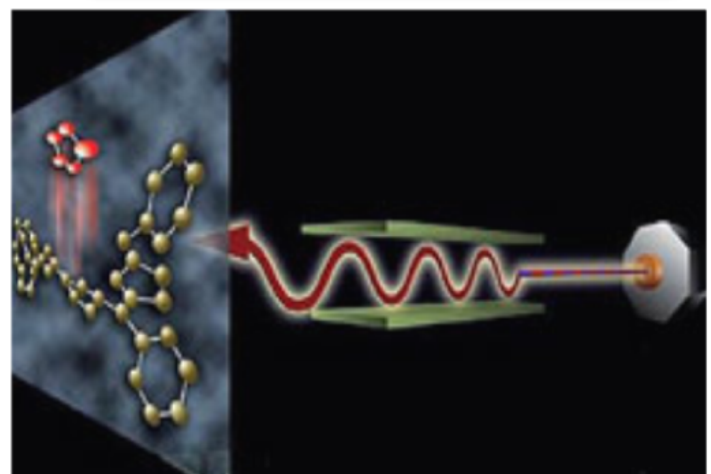


TeV e^-/e^+ beams for High energy physics



LHC

10s GeV e^- beams for Photon science (e.g., SLAC LCLS (Linac Coherent Light Source)

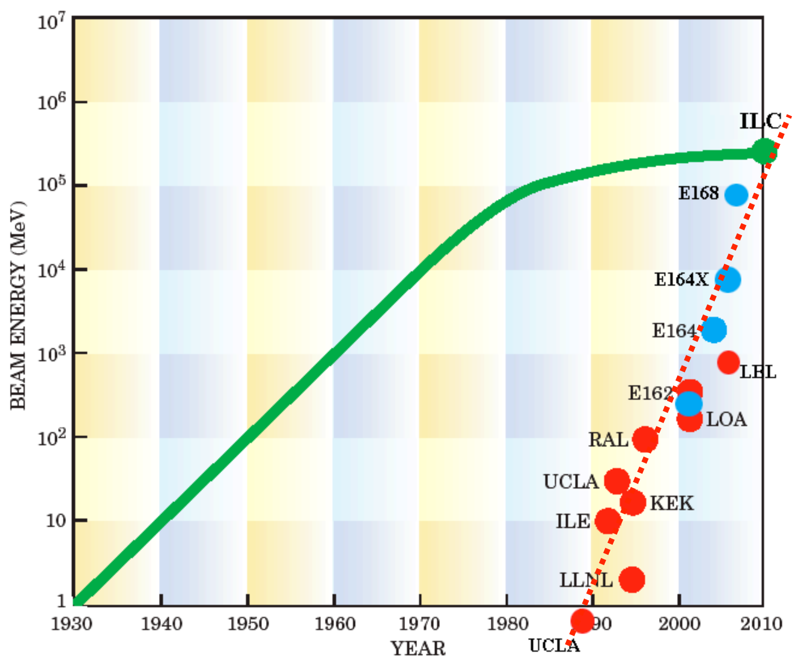


LCLS

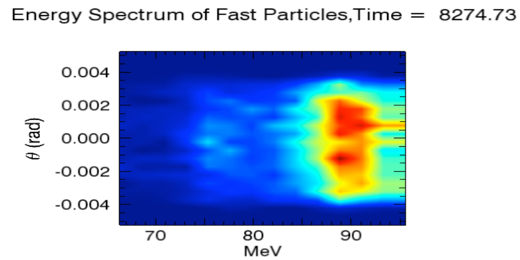
Livingston Curve for Accelerators --- Why plasmas?

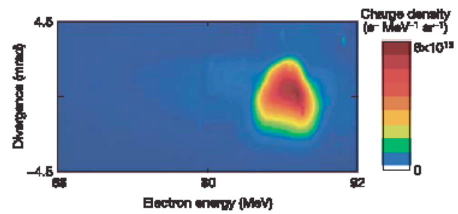
The Livingston curve traces the history of electron accelerators from Lawrence's cyclotron to present day technology. The curve clearly shows that conventional technology has reached saturation. The reason for this saturation is that conventional accelerators uses metal which has a "breakdown limit" and has a upper limit for how much electric field it can support, so because energy = force x distance, the only way to increase energy gain has been to build larger and larger accelerators.

When energies from plasma based accelerators are plotted in the same curve, it shows the exciting trend that within a few years it is will surpass conventional accelerators in terms of energy.

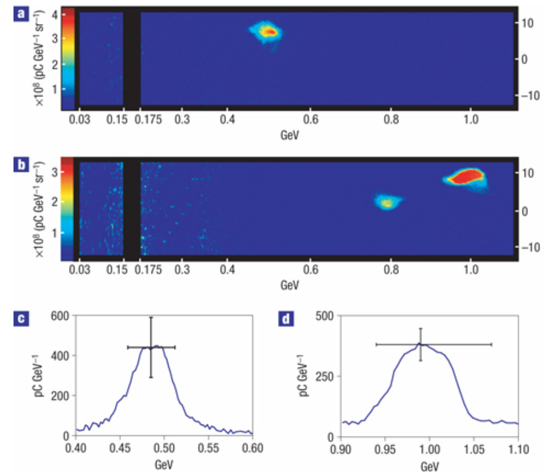


Recent Highlights in Plasma Based Acceleration (< Last 5 years) -- Simulations play a big role in all of these discoveries!!!

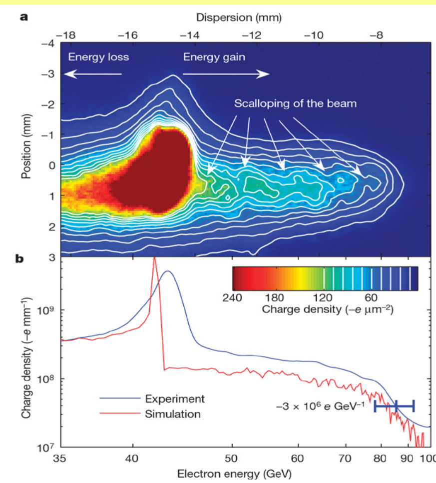


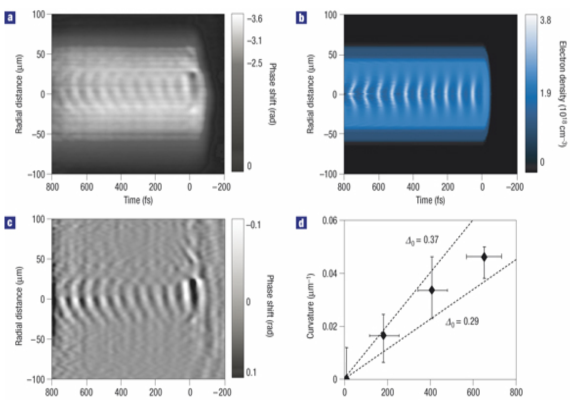


"Dream Beam" (Nature, 2004) -- 3 groups observed monoenergetic bunches using short (< 100fs) pulse lasers -- 3D simulations produced qnantitative agreements!!

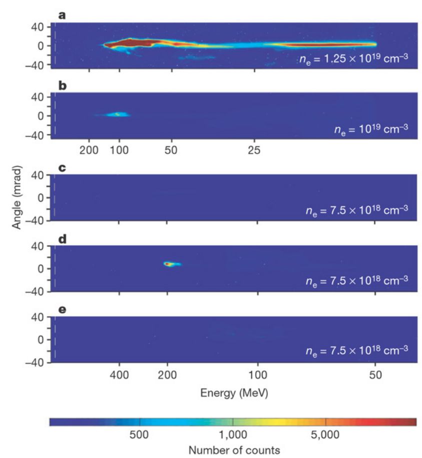


GeV LWFA in cm scale plasma





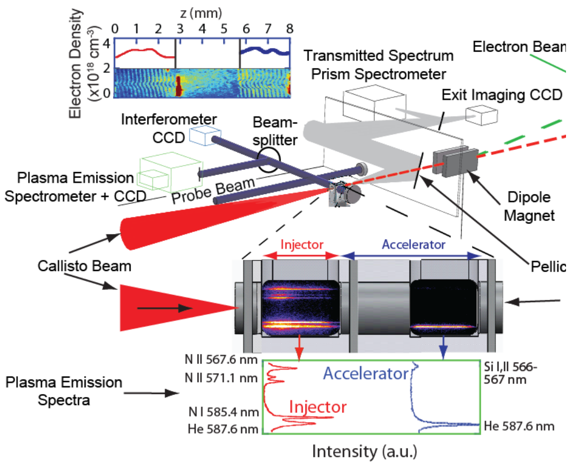
Laboratory snap shot of wakefield



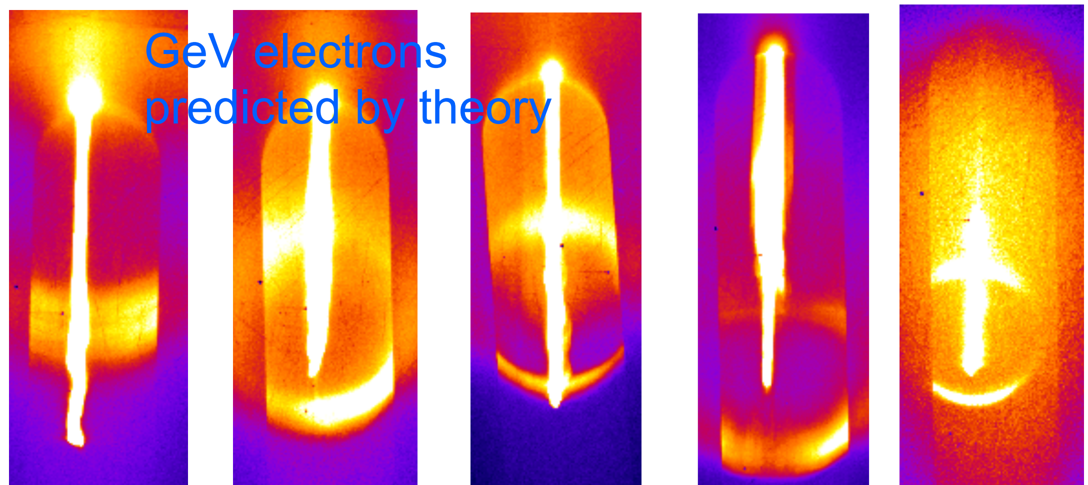
Controlled electron injection

42 GeV in less than one meter!
(i.e., 0-42 GeV in 3km, 42-85 GeV in 1m)
Simulations also identified ionization induced erosion as the limiting mechanism for energy gain

3D OSIRIS Simulations of Ring Formations in Callisto (LLNL) Experiments --- Every 20 shots or so a ring of electron is observed (the rings are very reproducible and have similar energy and angular spread)

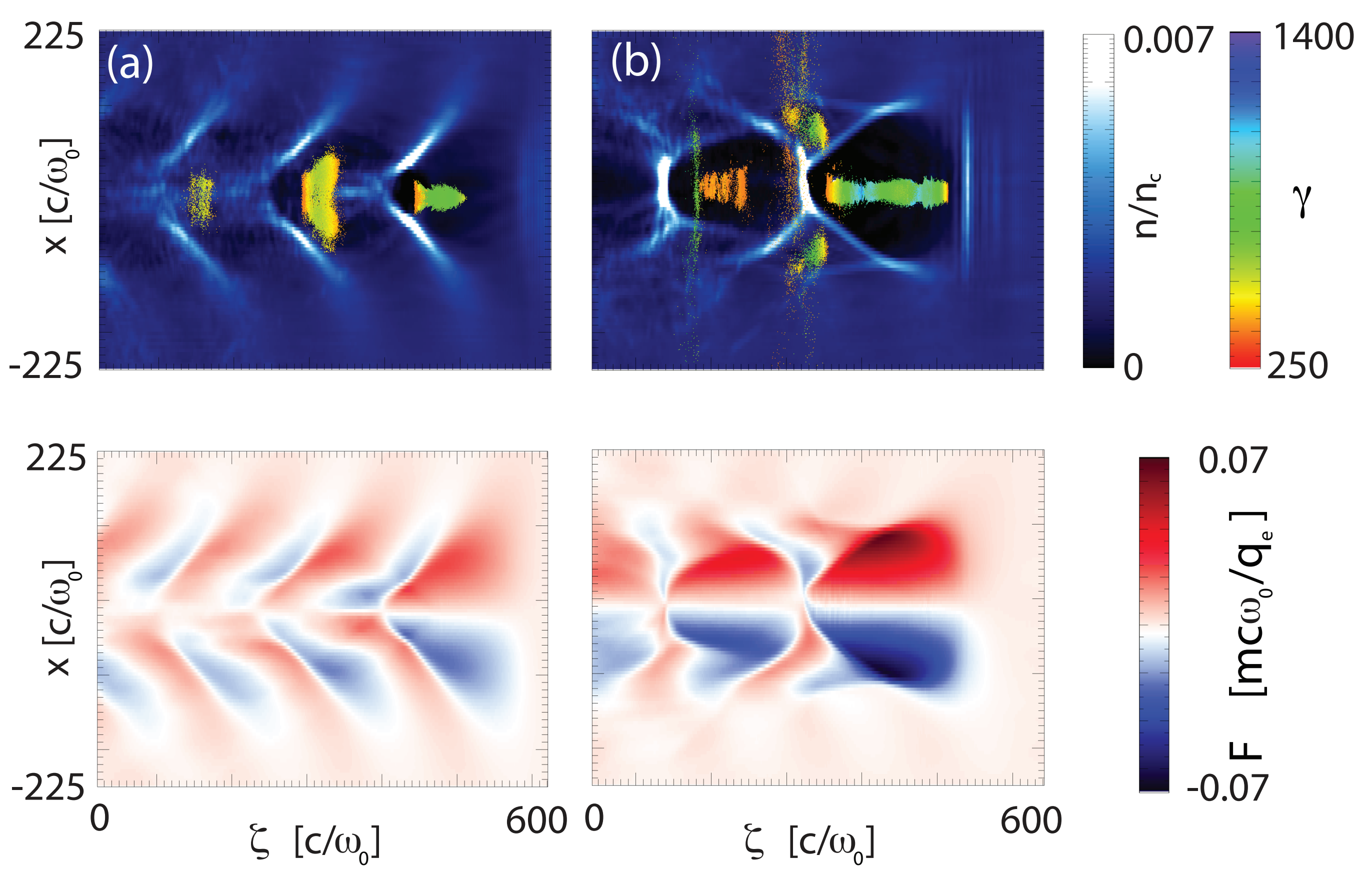


1309 | 1355 | 1360 | 1381 | 1399



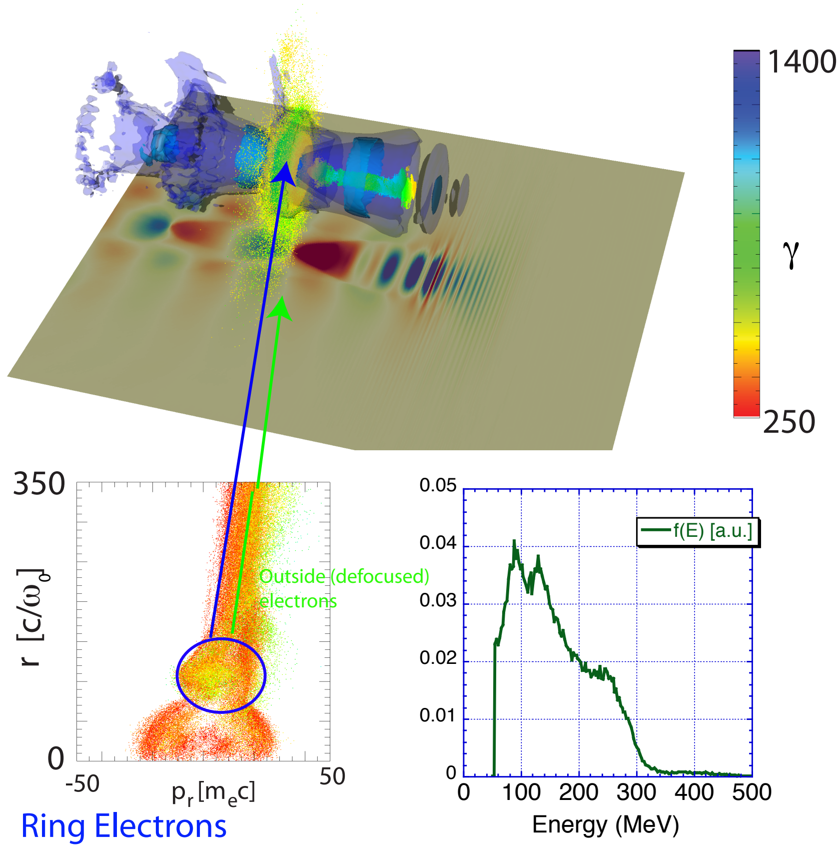
Length (mm)	6	10	10	10	10
Gas (He/N2)	100/0	90/10	99/1	98.5/1.5	100/0
Density (10^{18}/cc)	5	5	5	7	7
Power (TW)	77	108	113	128	113
Ring Energy (MeV)	136	215	170	280	249
dE/E (%)	28	19	5.9	25	4.8
Cone Angle (mrad)	105	65	55	78	49
Centroid Angle (mrad)	53	46	14	37	22
Tail Angle (mrad)	34	41	0	28	29

OSIRIS Simulations showed that laser evolution, not electron interactions, led to the formation of the rings.



The ring formation is observed in 3D simulations (see upper right plot) Originally it is thought that the ring is form when the electrons from the second bucket interacts with the electrons in the first bucket. However, simulations reveal that the evolution of the laser causes the system to evolve from a quasi-nonlinear regime to the blowout regime, creating a pocket of focusing structure in the center and some side lobes, and the electrons in the second bucket sit in these lobes and form the

Our 3D simulations yielded good quantitative agreements in beam energy, quality, and gave insights to its generation

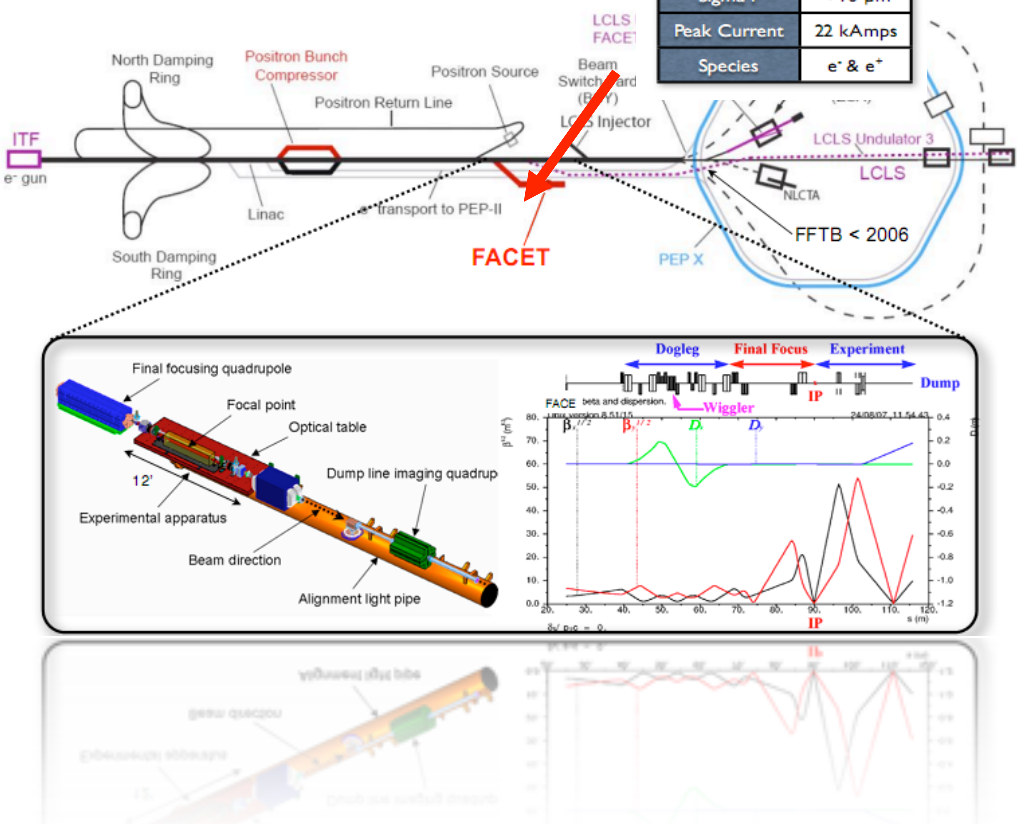


On the left you see the formation of the ring at the end of the simulation. (On the top figure, the plasma density makes up the isosurfaces and the energetic electrons are shown as dots.

3D simulations correctly reproduces the beam energy, the ring structure, and also the cone angle, furthermore, it provided insights into the mechanism which led to the ring formation.

QuickPIC Simulations of PWFA's (Dr. Weiming An)

Energy	24 GeV
Charge	3 nC
Sigma z	17 μm
Sigma r	< 10 μm
Peak Current	22 kAmps
Species	e^- & e^+



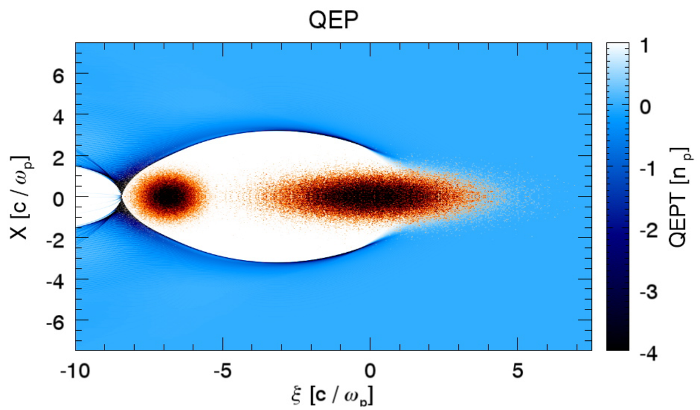
QuickPIC has been involved in the design of a linear collider using PWFA's using the SLAC beam (in the FACET facility, shown on the left) Some of the issues being studied are:

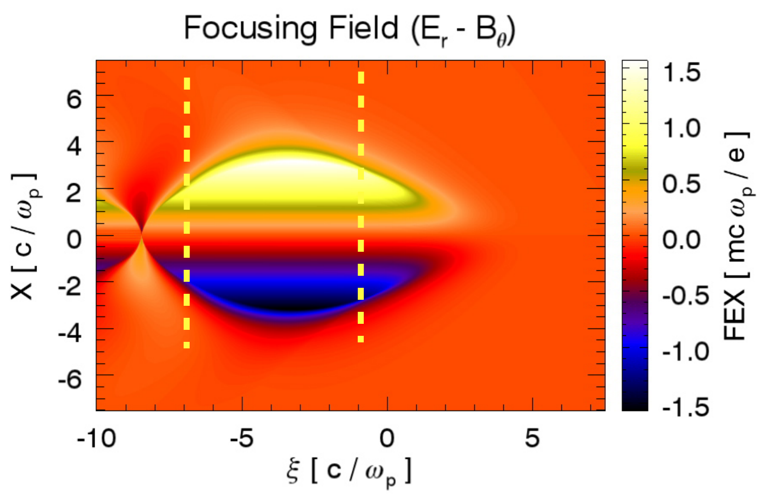
- Positron PWFA's (an accelerator for particle physics must have both particles and anti-particles, however, the theory for PWFA's using positron drivers is not well developed in the highly nonlinear regime because these beams do not propagate through anti-plasmas).
- Small energy spread (required to achieve luminosity and luminosity spectrum)
- **Small emittance (i.e., transverse profile)** and small emittance dilution (required to achieve luminosity). -- This requires very fine resolution in the transverse dimensions and is therefore memory intensive. Blue Waters is the ideal location to study this effect.
  - **e.g. Ion Motion**

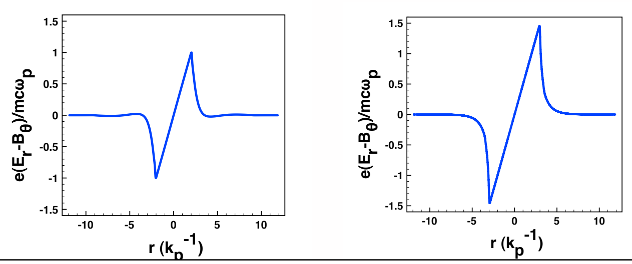
FACET is a new facility to provide high-energy, high peak current e^- & e^+ beams for PWFA experiments at SLAC, the goal is to achieve high efficiency, with low energy spread and low emittance. (In 2006 this facility demonstrates energy doubling in 1 meter using a long beam) Particle-in-cell simulations played a big role in understanding the detailed physics in that experiment.

Introduction to the blowout regime and ion motion

Blowout PWFA -- In a blowout PWFA, the driving bunch expels all of the electrons inside the accelerating structure, the ions inside the structure contains a linear focusing force, which preserves the *emittance* of the beam, which is the measure of the beam quality.







Ion Collapse

Electron beam

Electron beam

Ion rest frame

C frame

Flat-top beam

$$m_{ion}\frac{d^2r}{d\xi^2}+\frac{Zn_b}{2}r=0$$

$$\Delta\phi \approx k_{ion}\sigma_z \qquad k_{ion}=\sqrt{\frac{Zn_b}{2m_{ion}}}$$

$$\Delta\phi \approx \frac{k_{ion}}{k_p}k_p\sigma_z \approx \frac{k_{ion}}{k_p}=\sqrt{\frac{Zn_b}{2m_{ion}}}$$

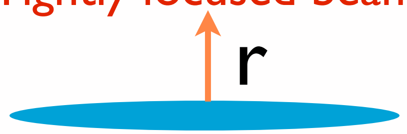
$$\boxed{\Delta\phi \approx \sqrt{\frac{Zn_b m_e}{2n_p m_{ion}}}} \qquad \boxed{\frac{n_b}{n_p} \gg \frac{m_{ion}}{m_e} \Rightarrow \Delta\phi \gg 1}$$

* J. B. Rosenzweig et al. Phys. Rev. Lett., 95:195002, 2005

The Plasma Ion Motion -- 3D QuickPIC Simulations

UCLA

Tightly focused beam



Flat-top beam

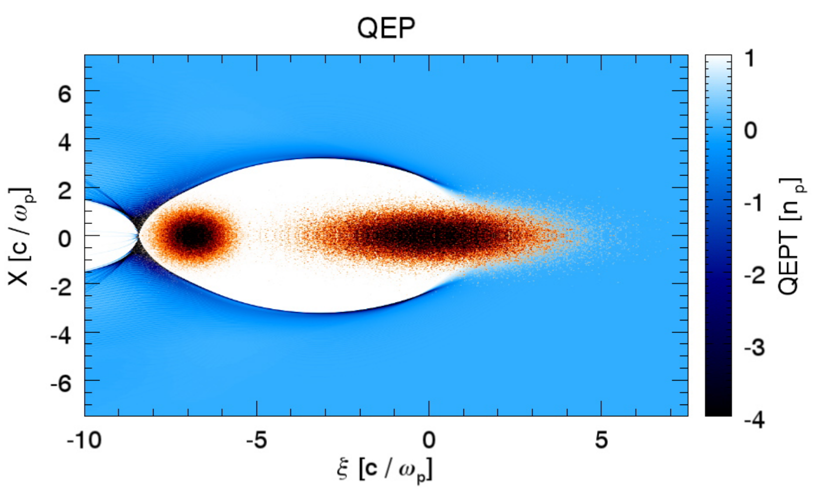
$$m_{ion}\frac{d^2r}{d\xi^2}+\frac{Zn_b}{2}r=0$$

$$\Delta\phi \approx k_{ion}\sigma_z \qquad k_{ion} = \sqrt{\frac{Zn_b}{2m_{ion}}} \;\rightarrow\; \Delta\phi \approx \frac{k_{ion}}{k_p}k_p\sigma_z \approx \frac{k_{ion}}{k_p} = \sqrt{\frac{Zn_b}{2m_{ion}}}$$

$$\boxed{\Delta\phi \approx \sqrt{\frac{Zn_b m_e}{2n_p m_{ion}}}}$$

σ_r = 0.1 μm, σ_z = 10.0 μm, N = 1.0 x 10^{10}
(based on proposed linear collider design)

n_b/n_p = 63500 >> m_{ion}/m_e = 12700

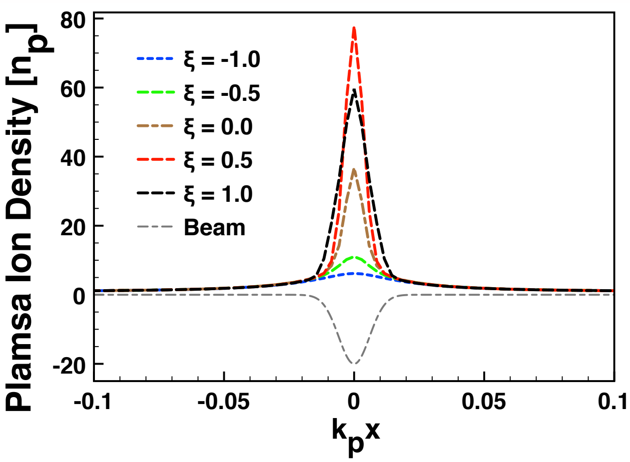


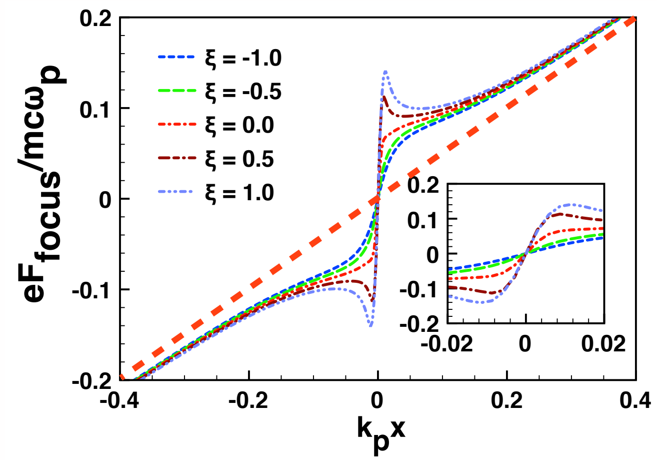
400 μm x 400 μm x 300 μm Box

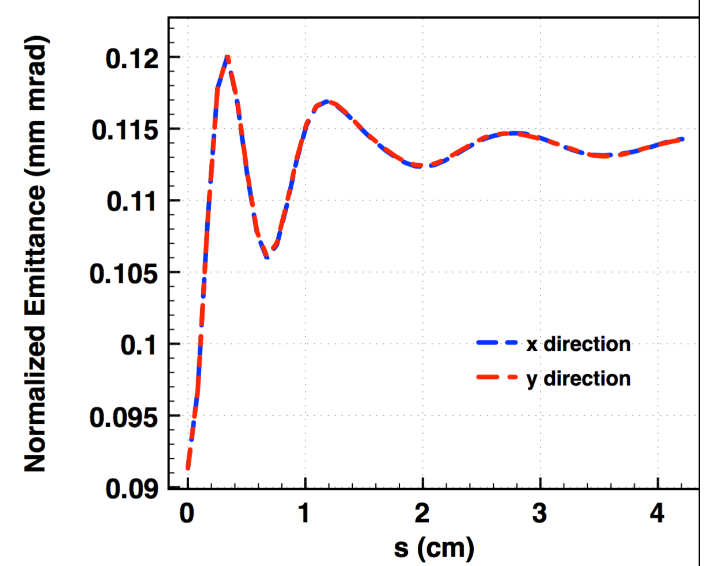
8000 x 8000 x 2000 Cells
550 million plasma particles
10 million beam particles (1 simulation electron = 1,000 real electrons)
typical simulation: > 32,000 cores, 1 million core hours (~5cm plasma propagation)

Ion Motion in PWFA

Emittance growth of the trailing beam



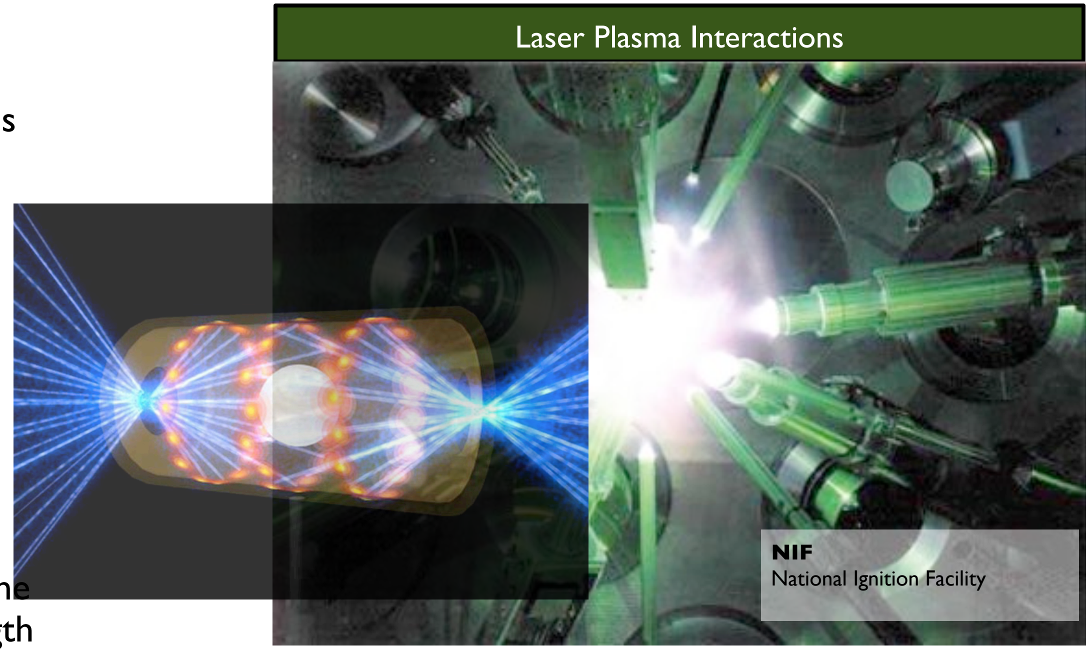




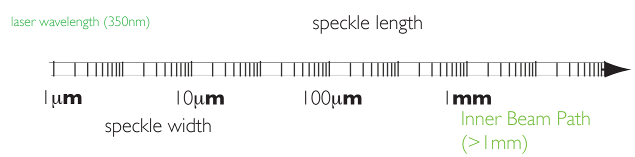
Simulations show that the trailing beam will reach a steady state after several cm propagation in the plasma. We find that for round beams that the **ion density enhancement is indeed by factors of 100, but that the emittance only grows by around 20%**. This is a very important result in the PWFA community and this is submitted for publication.

Laser Plasma Interactions in IFE

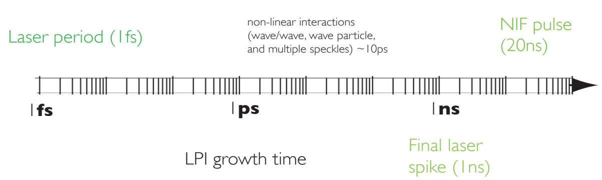
- Laser fusion uses lasers to compress fusion pellets. In this case, laser plasma interaction, where the incident laser decays into counter-propagating daughter waves. is important in IFE in 2 ways
  - LPI produces hot electrons which heats the target, making it harder to compress.
  - Laser light can be scattered backward toward the source and cannot reach the target
- The LPI problem is very challenging because of the various scales involved
  - The spatial scale spans from sub-micron (which is the laser wavelength) to mille-meters (which is the length of the plasma).
  - The temporal scale spans from a femto-second(which is the laser period) to nano-seconds (which is the duration of the fusion pulse)

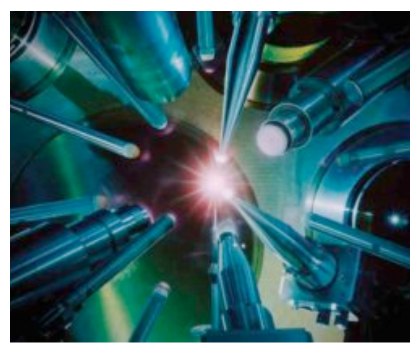


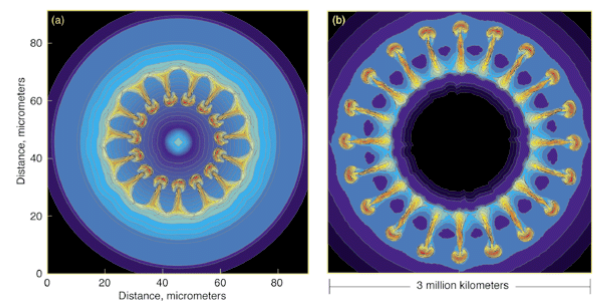
Lengthscales



Timescales







Striking similarities exist between hydrodynamic instabilities in (a) inertial confinement fusion capsule implosio and (b) core-collapse supernova explosions. [Image (a) is from Sakagami and Nishihara, *Physics of Fluids B* 2 2715 (1990); image (b) is from Hachisu et al., *Astrophysical Journal* **368**, L27 (1991).]

UCLA

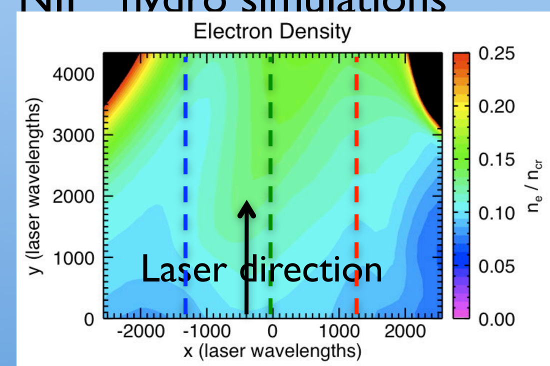
Currently most kinetic simulations of LPI's for NIF are done in 1D

- 1D simulations are quick and allow for methodical parameter scans and comparisons with linear theory. Currently, experimentalists @ NIF can re-construct plasma conditions (such as density and temperature) using a hydro code, and LPI information can be calculated using these plasma conditions.
  - Hydro conditions $\longrightarrow$ NIF uses 1D fluid postprocessing tools such as SLIP/NEWLIP:

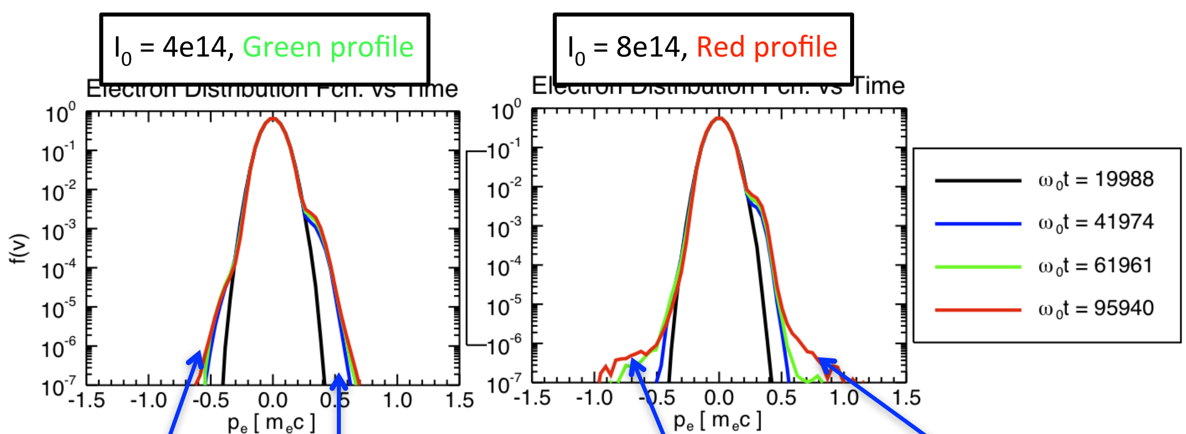
    Predict the frequency and reflectivity of the most unstable LPI
  - Hydro conditions $\longrightarrow$ 1D OSIRIS simulations:

    Similar capabilities + detailed information about energy partition (i.e., how much light reaches the target, how much light is reflected, and how much light is converted into hot electrons), backscattered light spectrum, and temperature of the hot electrons (show below).

I_{laser} = 2 – 8 x 10^{14} W/cm^2
λ_{laser} = 351nm,
T_e = 2.75 keV,
T_i = 1 keV, Z=1,
t_{max} up to 20 ps
Length = 1.5 mm
Density profiles from NIF hydro simulations



14 million particles
~400 CPU hours per run
~1 hr on modest size supercomputer



Due to backscatter

Due to LDI of backscatter

Due to LDI of rescatter

Due to rescatter of initial SRS

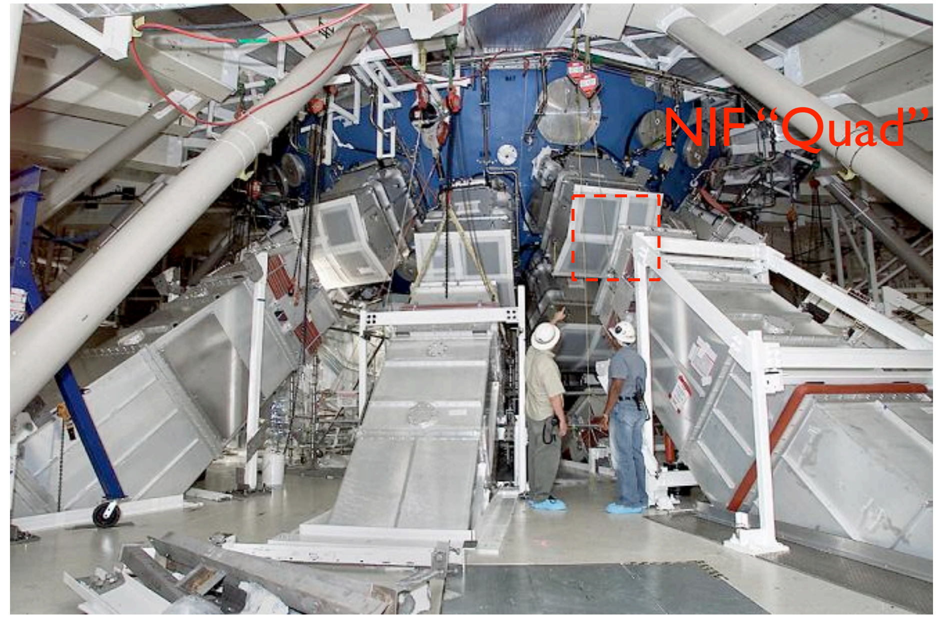
UCLA

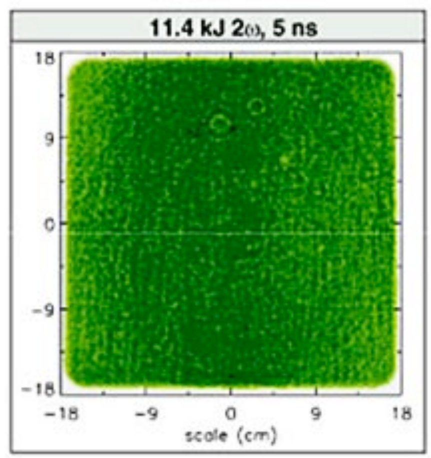
We have simulated stimulated Raman scattering in multi-speckle scenarios (in 2D)

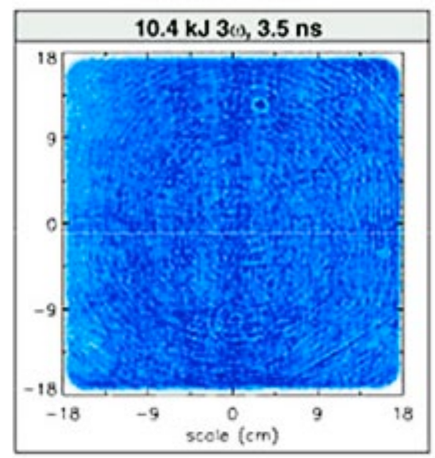
- Although the SRS problem is 1D (i.e., the instability grows along the direction of laser propagation). The SRS problem in IFE is not strictly 1D -- each "beam" (right) is made up of 4 lasers, called a NIF "quad," and each laser is not a plane wave but contains "speckles," each one a few microns in diameter. According to linear theory,The laser is LPI unstable only inside these "hotspots"

- We have been using OSIRIS to look at SRS in multi-speckle scenarios. In our simulations we observed the excitation of SRS in below-threshold speckles via:
  - "seeding" from backscatter light from neighboring speckles
  - "seeding" from plasma wave seeds from a neighboring speckle.
  - "inflation" where hot electrons from a neighboring speckle flatten the distribution function and reduce plasma wave damping. This can also trigger SRS

- The interaction of multiple speckles is a highly complex process and is ideally suited for PIC simulations







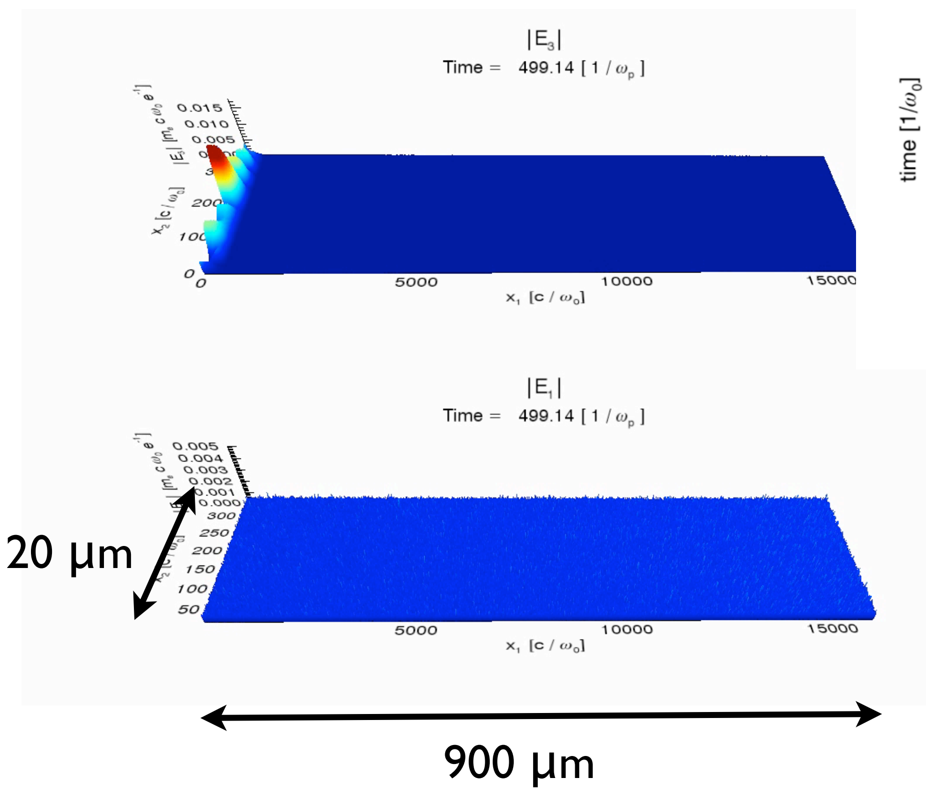
Multispeckle SRS Simulation using NIF density profiles

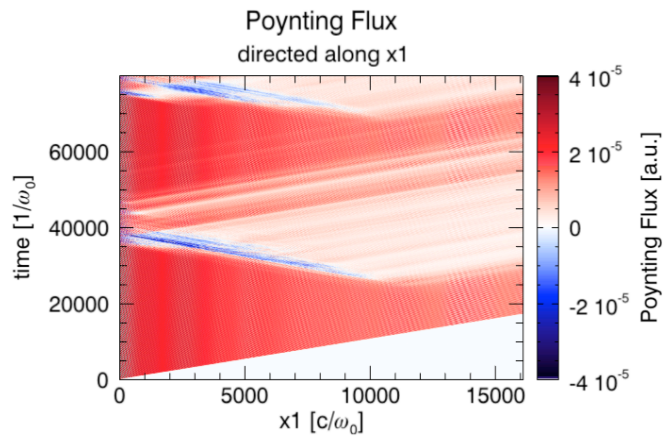
Speckled Laser Beam:

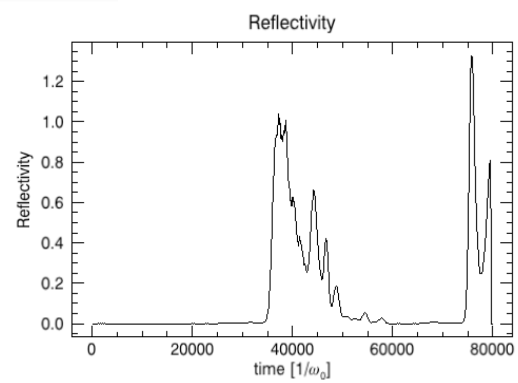
λ = 351nm

I_{avg} = 10^{15} W/cm^2

5 speckles long x 7 speckles wide







Domain:

6.4 million cells

16 billion particles

450,000 time steps

Computational:

32768 processors on Blue Waters

880,000 CPU hours

PIC simulations of LPI's is still a challenge, and requires exa-scale supercomputers, this will require code developments.

	2D multi-speckle along NIF beam path	3D, 2 speckles	3D, multi-speckle along NIF beam path
Speckle scale	50 x 8	2 x 1	10 x 10 x 5
Size (microns)	150 x 1500	18 x 9 x 120	28 x 28 x 900
Grids	9,000 x 134,000	1,000 x 500 x 11,000	1,700 x 1,700 x 80,000
Particles	300 billion	620 billion	22 trillion
Steps	470,000 (15 ps)	180,000 (5 ps)	540,000 (15 ps)
Memory Usage*	7 TB	18 TB	1.6 PB
CPU-Hours	8 million	9 million	1 billion (2 months on the full BW)

*memory usage can be reduced by the use of higher order particle shapes to reduce noise

Designing New Particle-in-Cell (PIC) Algorithms on GPU's -- add a new layer of parallelism on the processor.

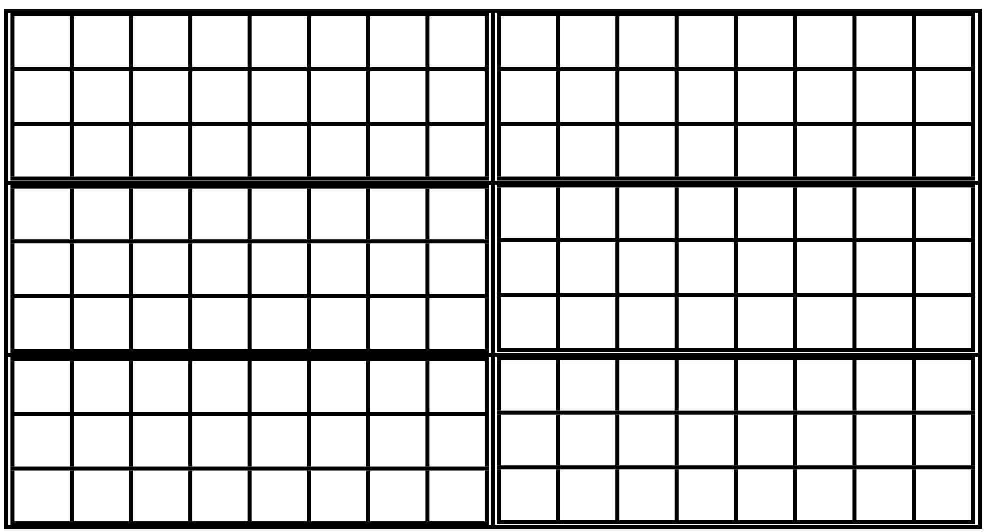
Particles ordered by tiles, varying from 2 x 2 to 16 x 16 grid points
On Fermi M2090:

- Associate a **thread block** with each tile and particles located in that tile

We created a new data structure for particles, partitioned among threads blocks (i.e., particles are sorted according to its tile id, and there is a local domain decomposition within the GPU), within the tile the particles and the particle data are aligned and the loops can be easily parallelized inside the GPU

We created a new data structure for particles, partitioned among threads blocks:

```
dimension part(npmax,idimp,num_blocks)
```

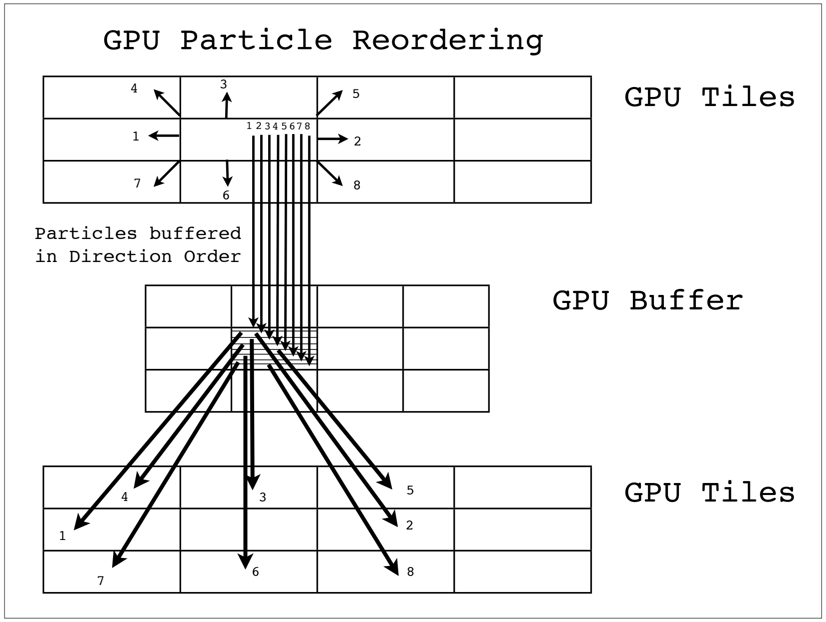


Designing New Particle-in-Cell (PIC) Algorithms:
Maintaining Particle Order

Three steps:
1. Particle Push creates a list of particles which are leaving a tile
2. Using list, each thread places outgoing particles into an ordered buffer it controls
3. Using lists, each tile copies incoming particles from buffers into particle array

• Less than a full sort, **low overhead** if particles already in correct tile
• **Essentially message-passing**, except buffer contains multiple destinations

In the end, the particle array belonging to a tile has no gaps
• Particles are moved to any existing holes created by departing particles
• If holes still remain, they are filled with particles from the end of the array



Evaluating New Particle-in-Cell (PIC) Algorithms on GPU: **Electromagnetic Case**
2-1/2D EM Benchmark with 2048x2048 grid, 150,994,944 particles, 36 particles/cell
optimal block size = 128, optimal tile size = 16x16

GPU algorithm also implemented in OpenMP

Hot Plasma results with dt = 0.04, c/vth = 10, relativistic

	CPU:Intel i7	GPU:Fermi M2090	OpenMP(12 CPUs)
Push	66.5 ns.	0.426 ns.	5.645 ns.
Deposit	36.7 ns.	0.918 ns.	3.362 ns.
Reorder	0.4 ns.	0.698 ns.	0.056 ns.
Total Particle	103.6 ns.	2.042 ns.	**9.062 ns (11.4x speedup)**.

The time reported is per particle/time step.
The total particle speedup on the Fermi M2090 was 51x compared to 1 Intel i7 core.

Field solver takes an additional 10% on GPU, 11% on CPU.

OK, so how about multiple CPU/GPU's?

Evaluating New Particle-in-Cell (PIC) Algorithms on GPU: **Electromagnetic Case**
2-1/2D EM Benchmark with 2048x2048 grid, 150,994,944 particles, 36 particles/cell
optimal block size = 128, optimal tile size = 16x16. Single precision. Fermi M2090 GPU

Hot Plasma results with dt = 0.04, c/vth = 10, relativistic

	CPU:Intel i7	1 GPU	2 GPUs	3 GPUs
Push	66.5 ns.	0.422 ns.	0.211 ns.	0.141 ns.
Deposit	36.7 ns.	0.972 ns.	0.488 ns.	0.327 ns.
Reorder	0.4 ns.	0.690 ns.	0.346 ns.	0.232 ns.
Total Particle	103.6 ns.	2.092 ns.	1.200 ns.	0.726 ns.

The time reported is per particle/time step.
The total speedup on the 3 Fermi M2090s compared to 1 core was 142x,
Speedup on 3 M2090s compared to 1 M2090 was 2.9x

Field solver takes an additional 10% on 1 GPU, 27% on 2 GPUs, and 52% on 3 GPUs (due to FFT's) --- strong scaling flattens out @ around 100 GPU's @ the UCLA Dawson2 cluster.

PIC Algorithms on future architectures are largely a hybrid combination of previous techniques

- Vector techniques from vector Cray's
- Blocking techniques from cache-based architectures
- Message-passing techniques from distributed memory architectures

Scheme should be portable to other architectures with similar hardware abstractions (such as the intel Phi)

Further information available at:
http://www.idre.ucla.edu/hpc/research/

Source codes available at:
https://idre.ucla.edu/hpc/parallel-plasma-pic-codes/ and the UCLA PICKSC web-site

UCLA Particle-in-Cell and Kinetic Simulation Software Center (PICKSC), NSF funded
Goal is to provide and document parallel Particle-in-Cell (PIC) and kinetic codes.

Planned activities

- Provide parallel skeleton codes for various PIC models on traditional and new parallel hardware and software systems.
- Provide MPI-based production PIC codes that will run on desktop computers, mid-size clusters, and the largest parallel computers in the world.
- Provide key components for constructing new parallel production PIC codes for electrostatic, electromagnetic, and other codes.
- Provide interactive codes for teaching of important and difficult plasma physics concepts
- Facilitate benchmarking of kinetic codes by the physics community, not only for performance, but also to compare the physics approximations used
- Documentation of best and worst practices, which are often unpublished and get repeatedly rediscovered.
- Provide some services for customizing software for specific purposes

Key components and codes will be made available through standard open source licenses and as an open-source community resource, contributions from others are welcome.

And we are hiring good post-docs! (please contact me or Prof. Warren Mori, the PI of this project)