On the road toward real-time virtual prototyping of particle accelerators

Jean-Luc Vay

In collaboration with: C. Benedetti, A. Bhagatwala, S. Bulanov, E. Esarey, C. Geddes, D. Grote, P. Lee, W. Leemans, R. Lehe, M. Lobet, B. Loring, C. Mitchell, G. Penn, O. Ruebel, J. Qiang, C. Schroeder, C. Sun, H. Vincenti

March 26, 2016

Particle accelerators are essential tools in modern life that power scientific discovery, cure cancer, secure our borders, and help create a wide range of products

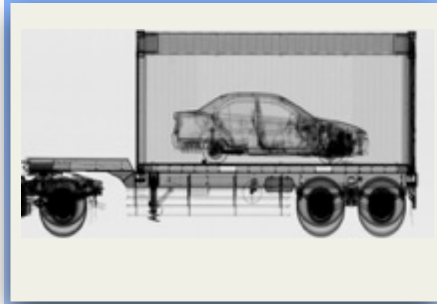
Medicine

- ~9000 medical accelerators in operation worldwide
- 10’s of millions of patients treated/yr
- 50 medical isotopes, routinely produced with accelerators

Industry

- ~20,000 industrial accelerators in use
  - Semiconductor manufacturing
  - cross-linking/ polymerization
  - Sterilization/ irradiation
  - Welding/cutting
- Annual value of all products that use accel. Tech.: $500B

National Security



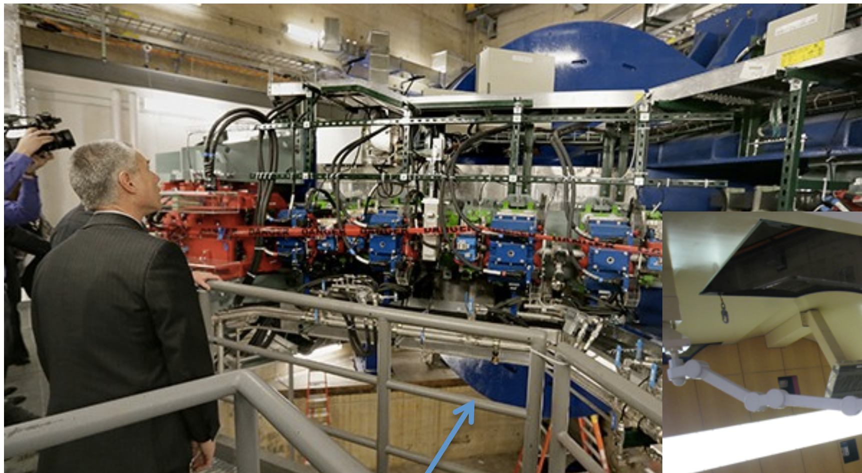
- Cargo scanning
- Active interrogation
- Stockpile stewardship: materials characterization, radiography, support of non-proliferation

Discovery Science

- ~30% of Nobel Prizes in Physics since 1939 enabled by accelerators
- 4 of last 14 Nobel Prizes in Chemistry for research utilizing accelerator facilities

There are 30,000 Particle Accelerators Making an Impact on Our Lives

Problem: size & cost often a limiting factor
Example 1: Proton Therapy Center



New Rochester Mayo Clinic
Proton Therapy Center

- 4 chambers
- $188M

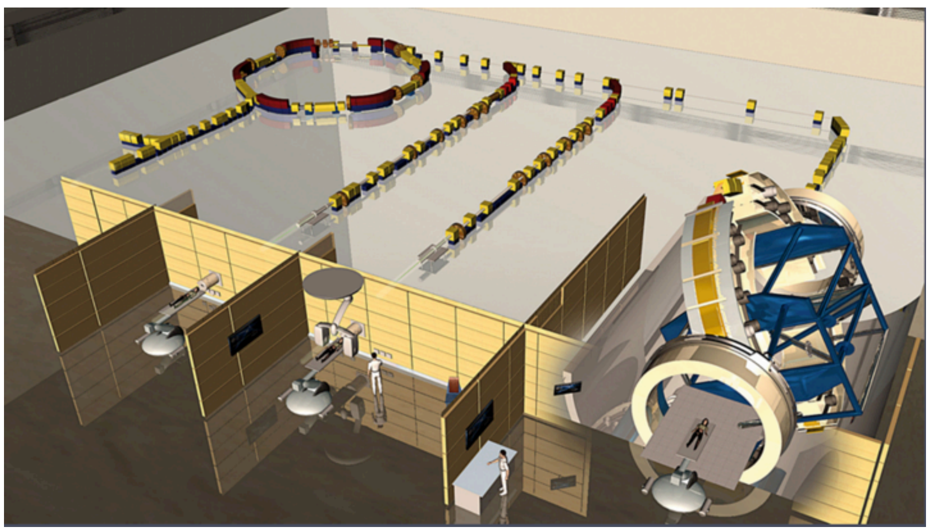
120-ton gantry directs proton beam
to appropriate spot on patient
by rotating around a three-story chamber.

http://finance-commerce.com/2014/03/status-report-mayo-proton-therapy-facility/#ixzz43DJgnIIA
http://blogs.mprnews.org/statewide/2014/03/mayos-proton-beam-facility-on-track-for-2015-opening/

Problem: size & cost often a limiting factor
Example 2: Carbon Therapy Center

Heidelberg Proton & Carbon Therapy Center
- 2 scans chambers
- one 4π chamber
- €119M

670-ton gantry
~1/10 Eiffel Tower

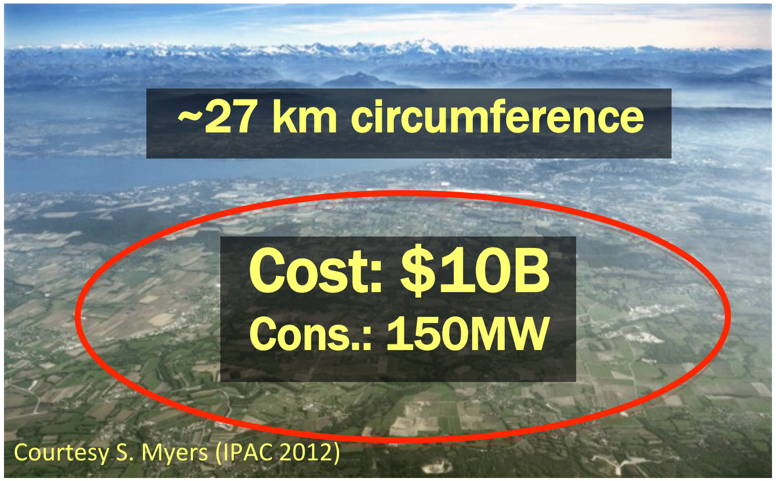


http://medicalphysicsweb.org/cws/article/research/51684
https://www.klinikum.uni-heidelberg.de/About-us.124447.0.html?&L=1

Problem: size & cost often a limiting factor

Example 3: High-Energy Physics collider

CERN LHC



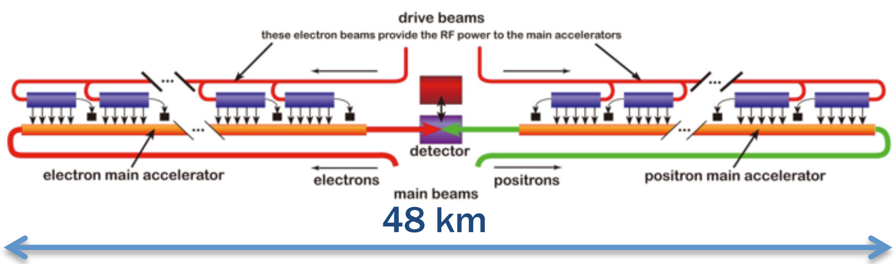
Future colliders?

ILC

Cost: $8B-$20B?

Cons.: 230MW

CLIC



Cost: $?B

Cons.: 415MW

FCC

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Au

News & Comment > News > Q&A > 2015 > December > Article

NATURE | NEWS: Q&A

CERN's next director-general on the LHC and her hopes for international particle physics

Fabiola Gianotti talks to *Nature* ahead of taking the helm at Europe's particle-physics laboratory on 1 January.

Elizabeth Gibney

22 December 2015

Fabiola Gianotti is the incoming director-general of CERN.

Some people think that future governments will be unwilling to fund larger and more expensive facilities. Do you think a collider bigger than the LHC will ever be built? And will it depend on the LHC finding something new?

The outstanding questions in physics are important and complex and difficult, and they require the deployment of all the approaches the discipline has developed, from high-energy colliders to precision experiments and cosmic surveys. High-energy accelerators have been our most powerful tools of exploration in particle physics, so we cannot abandon them. What we have to do is push the research and development in accelerator technology, so that we will be able to reach higher energy with compact accelerators.

Most applications of accelerators make use of conventional technology (with roots in post-WWII era)

This technology has been refined and extended...

1956: Stanford Medical Center

Today: Varian Medical

1940: Cockcroft Walton generator Cavendish Lab

Today: Superconducting resonators

Today, we are on the brink of technological breakthroughs...

Two-beam acceleration
—
Power transfer from one beam to another with RF structures

Dielectric Wakefield Accelerator
—
Strong fields induced in dielectric structure

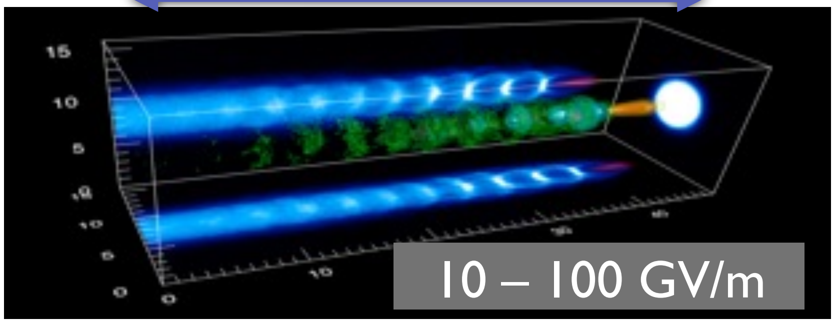
Beam-Plasma Wakefield Accelerator
—
Beam generated plasma wake with very high fields

Laser-Plasma Accelerator
—
Laser generated plasma wake with very high fields

Laser plasma acceleration enables development of compact accelerators

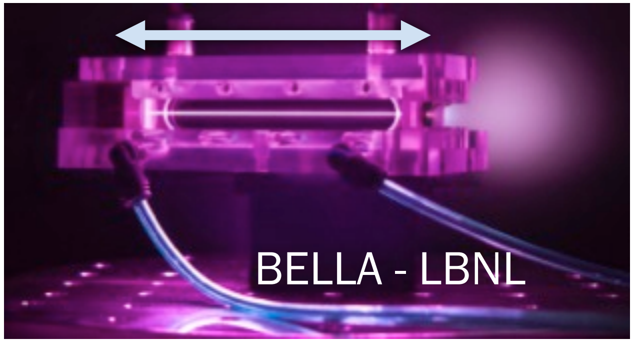
meter-scale

100 micron-scale

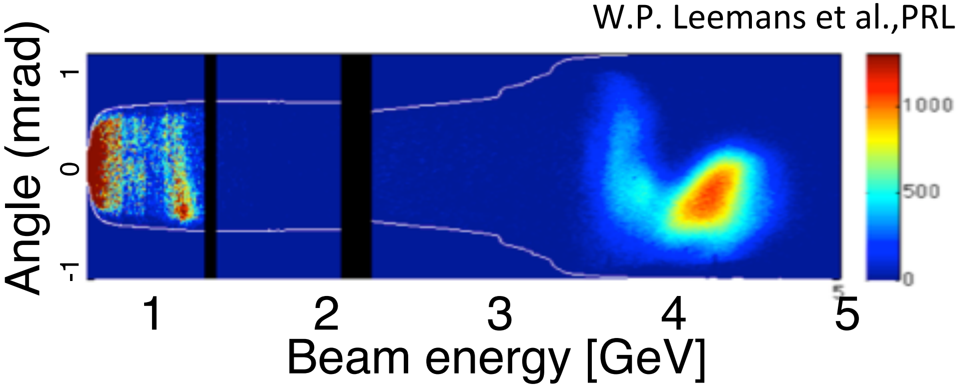


State-of-the-art

9 cm



W.P. Leemans et al.,PRL 2014

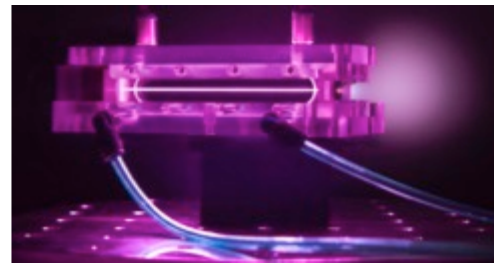


Multi-GeV electron beam from 9 cm laser plasma accelerator

Goal

Let's go from this...

...to this...



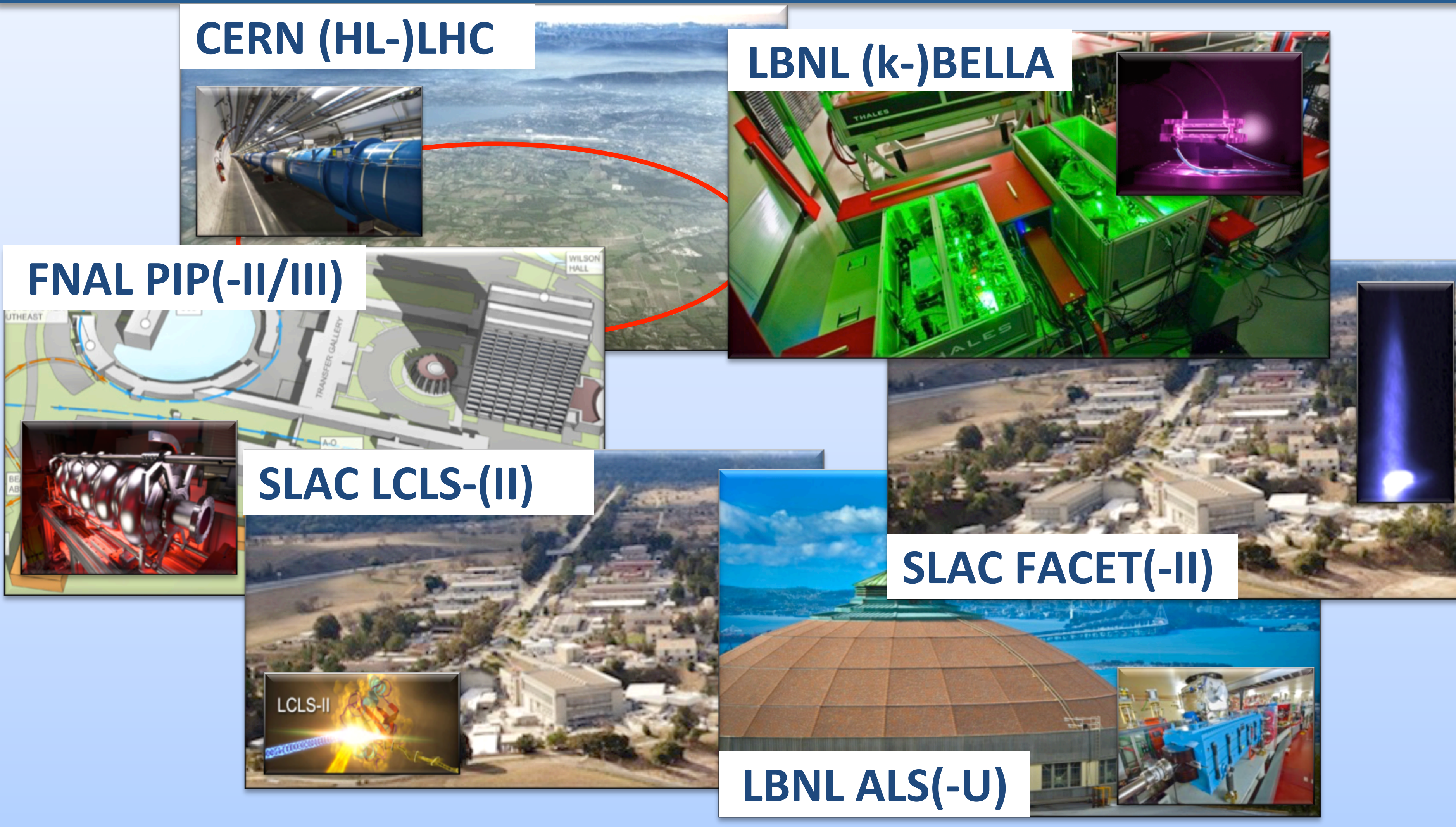
...and make it ready for application.

Can this technology be developed for light sources, medical and security applications, and even high energy colliders?

High-performance computing is key to answering these questions and reaching that goal!

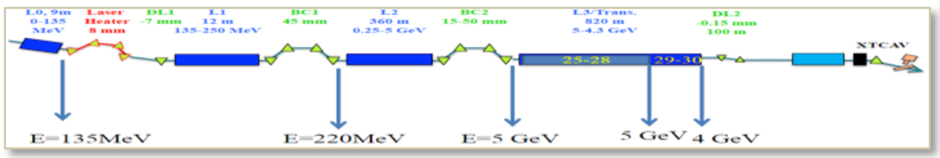
Modeling of particle accelerators

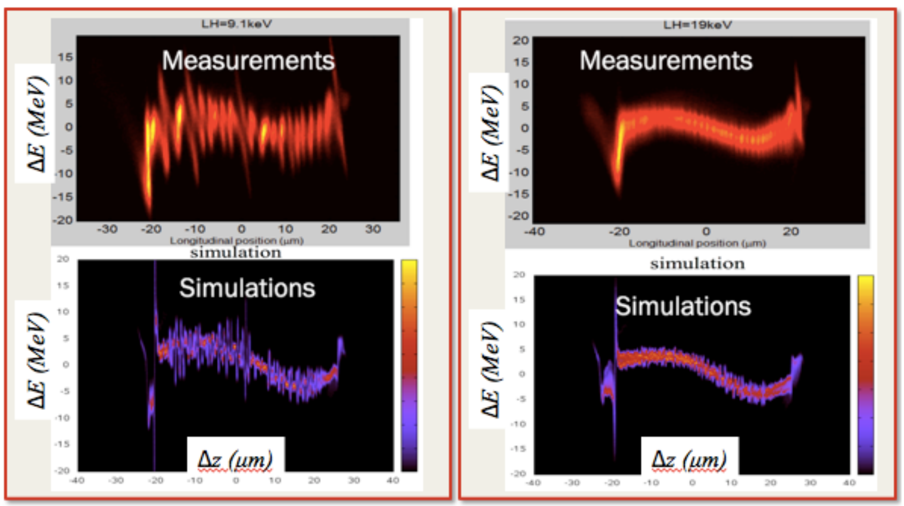
All accelerators in the world rely on modeling



Advanced simulations have key impact on *the support & analysis of experiments*

Start-to-end simulations of LCLS match measurements

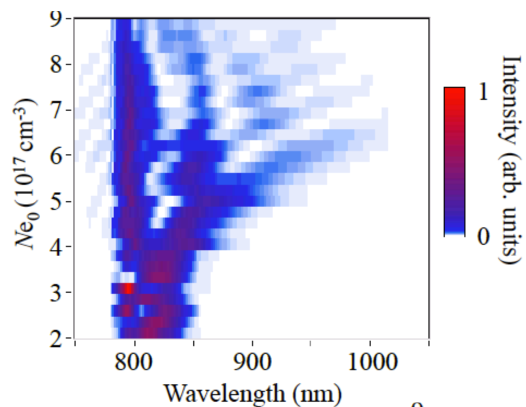




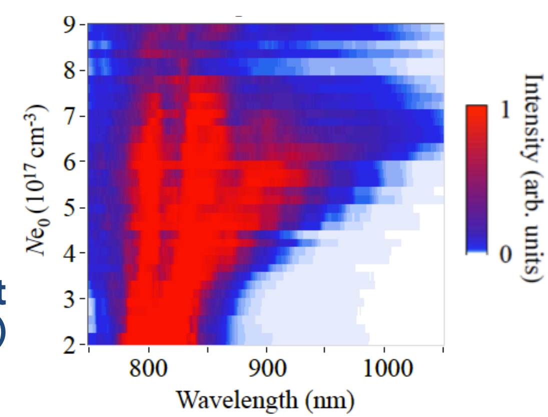
LH σ_E = 9.1 keV   LH σ_E = 19 keV

Simulations with IMPACT

Supported world record BELLA 4.25 GeV beam over 9 cm



Simulations (INF&RNO)



Experiment (BELLA)

*W. P. Leemans, et al., *Phys. Rev. Lett.* 113, 245002 (2014)

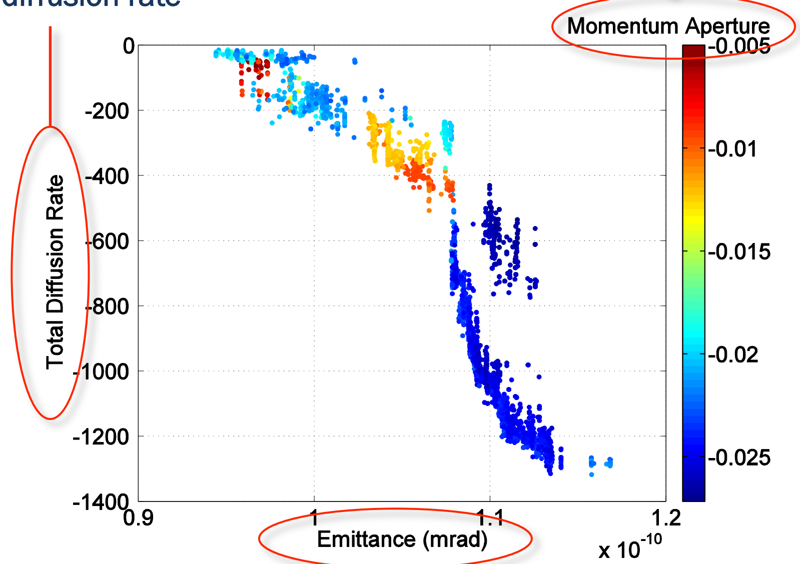
Advanced simulations also key to *the design of future accelerators*

Parallel genetic optimization improves design of ALS upgrade.

3 Objectives:

Momentum aperture

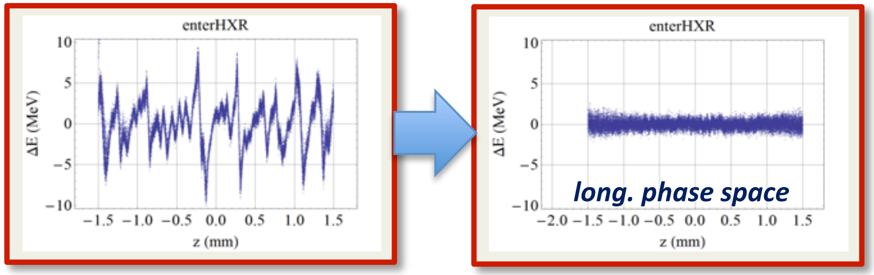
Total diffusion rate



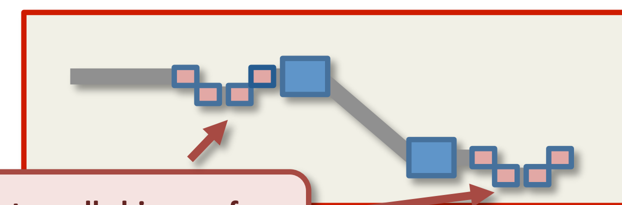
horizontal emittance

C. Sun *et al.*, (2015).

Predict amplification of instability in LCLS-II.



and validate novel mitigation scheme



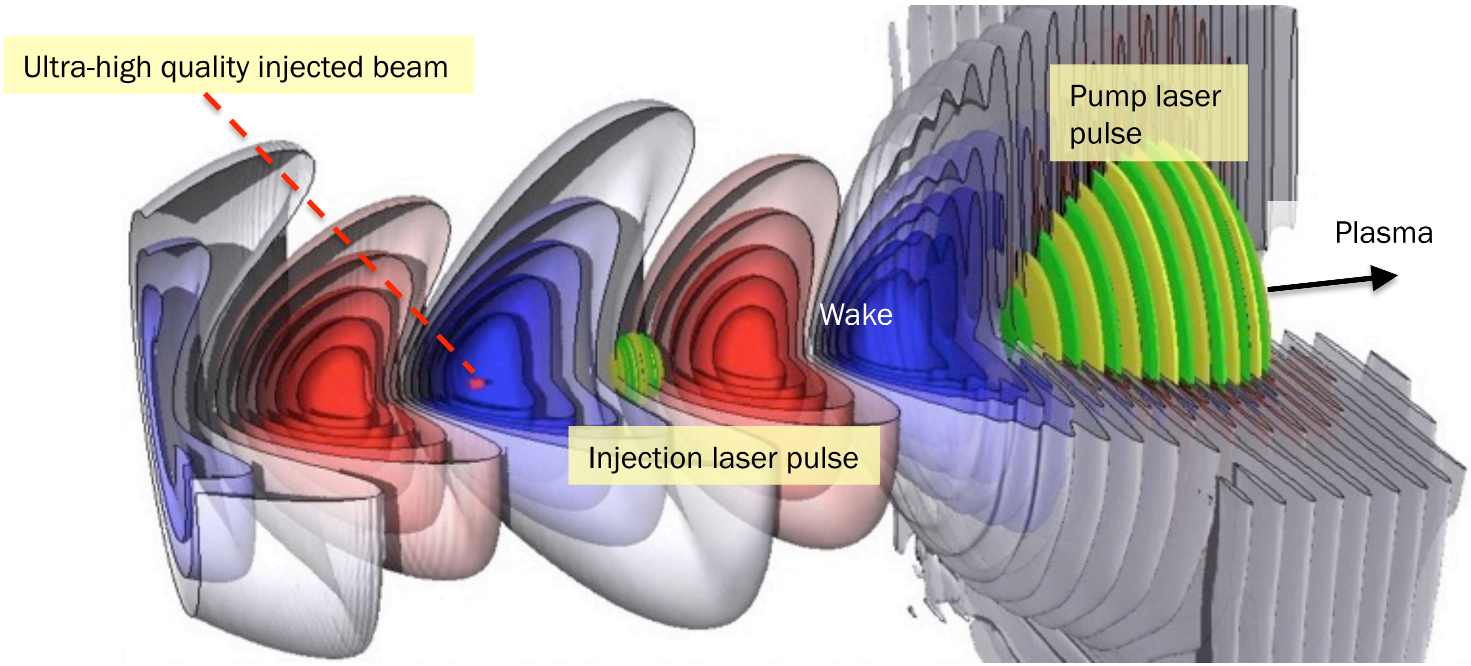
Insert small chicanes for local compensation of R_{56}

Simulations with IMPACT

M. Venturini and J. Qiang, PRST-AB 18, 054401 (2015).

Advanced simulations are essential tools for *the exploration of new concepts*

Demonstrate concept of two-color injection of ultra-high quality beam



Simulations with Warp - Visualization with VisIt

L.-L. Yu, et al, *Phys. Rev. Lett.* **112**, 125001 (2014)

Next generation of accelerators *needs* next generation of modeling tools

Our vision

Fast – *runs in seconds to minutes*

Hi-Fi – *full & accurate physics*

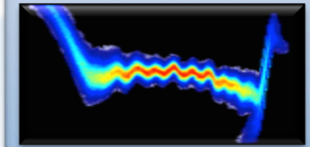
Link – *integrated ecosystem*

Real-time virtual prototyping of entire accelerator

with intuitive interface, dissemination & user support.

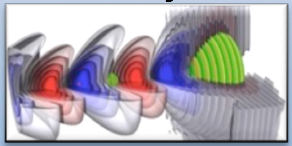
Simulations take too long!

X-FEL start-to-end



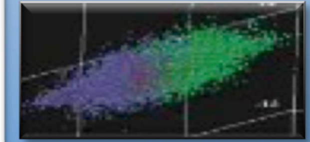
6 Hrs

2-color injection



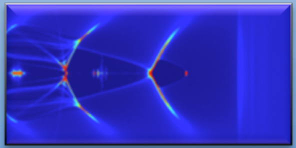
3 days

Beam-beam LHC



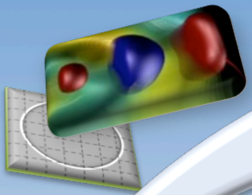
1 day

BELLA



7 days

Combine best algorithms



Speed

2016 Hours-days/run → 2025 Min./run

Port codes to fastest hardware

Need to speedup by ×10^n

Investment in new algorithms pushes the frontier

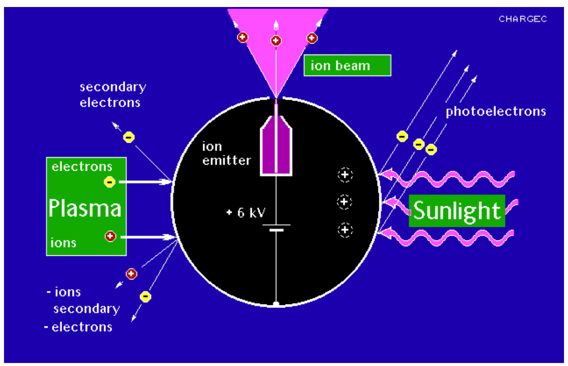
Algorithm/method	Reference	Originated	Adopted by
Integrated Maps for rf cavity dynamics	*Ryne, LANL Report* ***1995***	ML/IMPACT	D. Abell nonlinear model
Stochastic Leap-Frog for Brownian motion	*Qiang & Habib, PRE* ***2000***	IMPACT	
Spectral-finite difference multigrid solver	*Qiang & Ryne, CPC* ***2001/2006***	IMPACT	
Improved Perfectly Matched Layers	*Vay, JCP 2000/JCP* ***2002***	Warp	Osiris
AMR-PIC electrostatic	*Vay et al, LPB* ***2002****/PoP* ***2004***	Warp	
PIC w/ shift-Green function method	*Qiang et al, PRSTAB* ***2002****/CPC* ***2004***	BBeam3D	
Secondary emission of electrons algorithm	*Furman & Pivi, PRST-AB* ***2003***	Posinst	TxPhysics, Warp, spacecraft charging codes
AMR-PIC electromagnetic	*Vay et al, CPC* ***2004***	Emi2D	Warp
3D Poisson solver with large aspect ratio	*Qiang & Gluckstern, CPC* ***2004***	IMPACT	
PIC w/ integrated Green function	*Qiang et al, PRSTAB* ***2006***	ML/IMPACT	BB3D, Pyheadtail, Opal
Hybrid Lorentz particle pusher	*Cohen et al, NIMA* ***2007***	Warp	
Lorentz boosted frame 	*Vay, PRL* ***2007***	Warp	Osiris, Vorpal, PIConGPU, INF&RNO, JPIC, …

and it also benefits the entire community

Algorithm/method (cont.)	Reference	Originated	Adopted by
Explicit Lorentz invariant particle pusher	*Vay, PoP* ***2008***	Warp	Tristan, Osiris, PIConGPU, Photon-Plasma, QED, etc.
New convolution integral w/ smooth kernel	*Qiang, CPC* ***2010***	N/A	
Mixed Particle-Field decomposition method	*Qiang & Li, CPC* ***2010***	BBeam3D	
PIC with tunable electromagnetic solver	*Vay et al, JCP* ***2011***	Warp	Osiris, Vorpal
Efficient digital filter for PIC	*Vay et al, JCP* ***2011***	Warp	Osiris, Vorpal
Laser launcher from moving antenna	*Vay et al, PoP* ***2011***	Warp	Osiris, Vorpal
High-precision laser envelope model	*Benedetti et al,* ***2011***	Inf&rno	
Domain decomposition for EM spectral solver	*Vay et al, JCP* ***2013***	Warp	
Mitigation of num. Cherenkov instability	*Godfrey&Vay, JPC/CPC* ***2014-15***	Warp	Osiris
Adaptive unified differential evolution algo.	*Qiang & Mitchell, OO dig.* ***2015***		
Spectral solver with azimuthal decomposition	*Lehe et al, CPC, in press* ***2016***	FBPIC/Warp	**Latest with more in preparation**
Novel algorithm for vectorization	*Vincenti et al, submitted* ***2016***	Warp/PICSAR	
Generalized spectral solver	*Vay et al, in preparation* ***2016***	Warp	

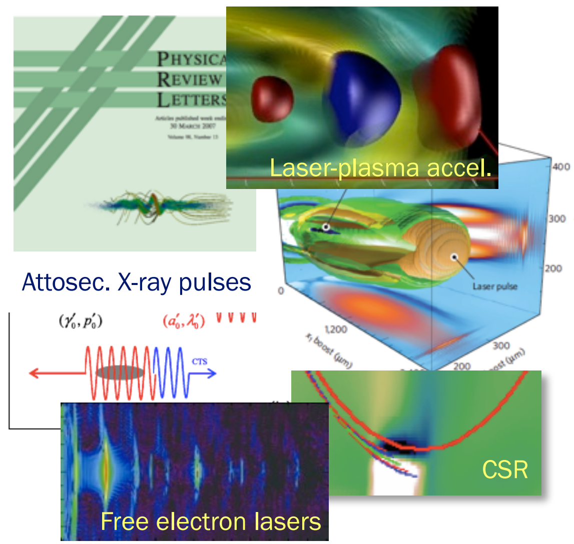
It also has relevance beyond original application

Secondary e- yield (SEY) model (Furman-Pivi, PRST-AB 2003)



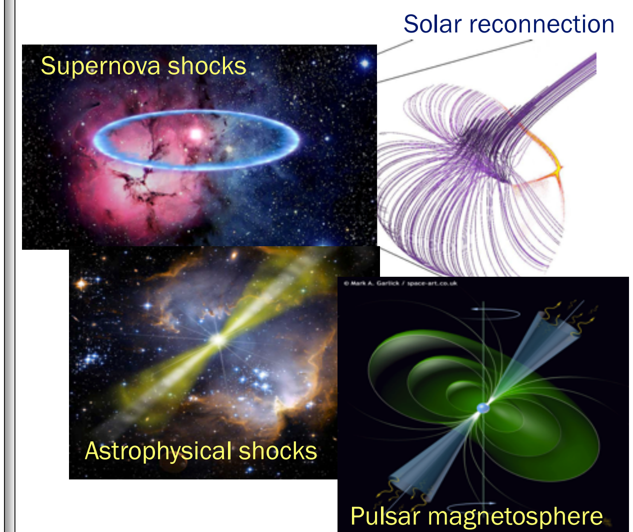
Prediction of spacecraft charging from charged particles impact.

Lorentz boosted frame (Vay, PRL 2007)



Speed-up first principles simulations x 10^n.

Relativistic particle pusher (Vay, PoP 2008)

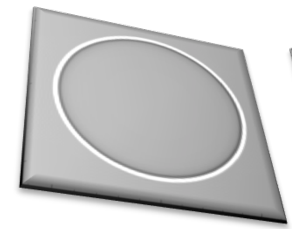


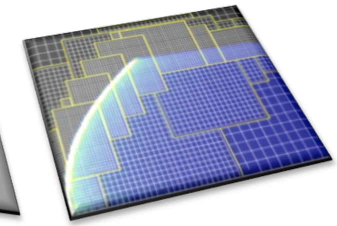
More precision integration of relativistic particle motion.

We are now combining algorithms for maximum efficiency

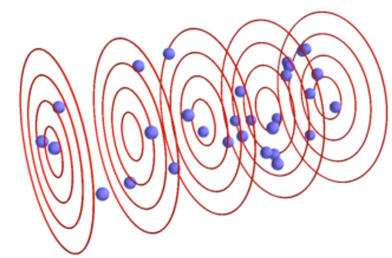
Lower # time steps:
- optimal Lorentz boosted frame

Higher accuracy:
- Lorentz invariant particle pusher
- Pseudo-spectral Maxwell solvers
- AMR

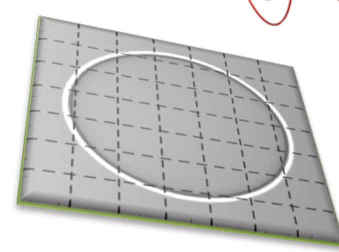




Lower dimensionality
- FFT+Hankel Transform Maxwell solver for quasi-RZ geom



Higher scalability
- FFT Maxwell solvers with domain decomposition



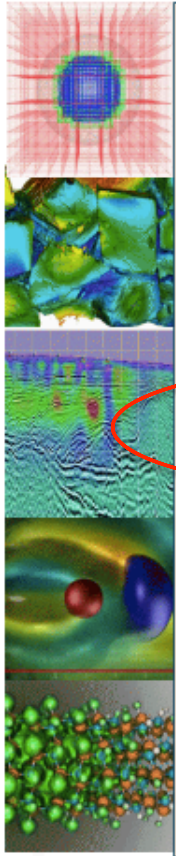
Higher stability
- Analysis & mitigation of Numerical Cherenkov Instability

Preparing Accelerator Codes for Exascale Supercomputing

Our proposal was selected as part of the NERSC Exascale Applications Program (NESAP)

NESAP Codes

NERSC



Advanced Scientific Computing Research

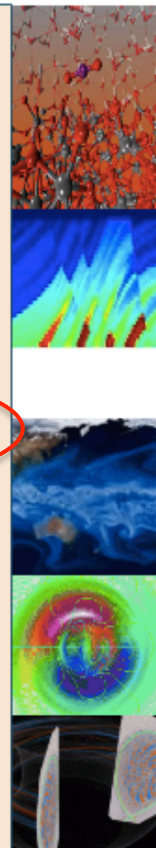
Almgren (LBNL)	**BoxLib AMR Framework**
Trebotich (LBNL)	**Chombo-crunch**

High Energy Physics

Vay (LBNL)	**WARP & IMPACT**
Toussaint(Arizona)	**MILC**
Habib (ANL)	**HACC**

Nuclear Physics

Maris (Iowa St.)	**MFDn**
Joo (JLAB)	**Chroma**
Christ/Karsch (Columbia/BNL)	**DWF/HISQ**



Basic Energy Sciences

Kent (ORNL)	**Quantum Espresso**
Deslippe (NERSC)	**BerkeleyGW**
Chelikowsky (UT)	**PARSEC**
Bylaska (PNNL)	**NWChem**
Newman (LBNL)	**EMGeo**

Biological and Environmental Research

Smith (ORNL)	**Gromacs**
Yelick (LBNL)	**Meraculous**
Ringler (LANL)	**MPAS-O**
Johansen (LBNL)	**ACME**
Dennis (NCAR)	**CESM**

Fusion Energy Sciences

Jardin (PPPL)	**M3D**
Chang (PPPL)	**XGC1**

U.S. DEPARTMENT OF ENERGY | Office of Science

Courtesy R. Gerber, NUGEX 2016

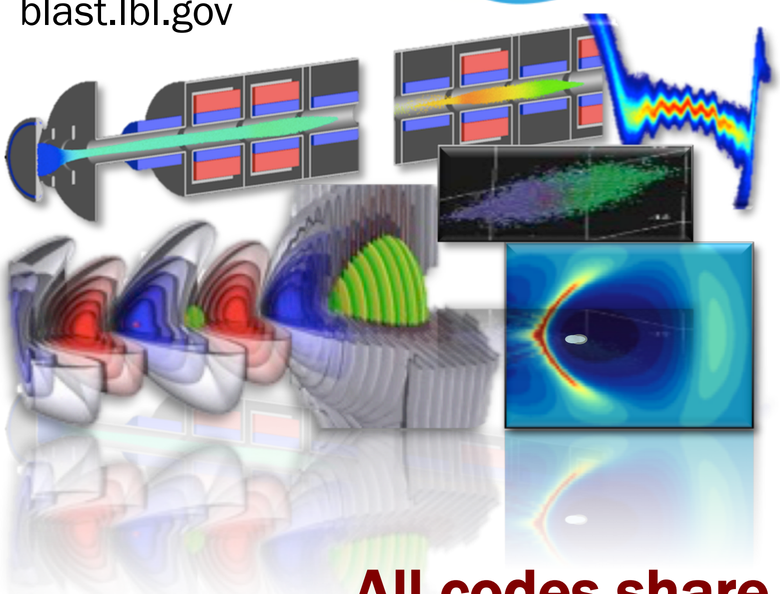
BERKELEY LAB

Mathieu Lobet

NESAP postdoc
Started on 2/1/16

LBNL home of unique simulation toolset
for conventional & advanced concepts accelerators

BLAST
BERKELEY LAB ACCELERATOR SIMULATION TOOLKIT
blast.lbl.gov



State-of-the-art codes:

- WARP, IMPACT, BEAMBEAM3D, INF&RNO, POSINST, FBPIC.

For detailed modeling of the largest set of physics and components:

- beams, plasmas, lasers, structures, etc *in* linacs, rings, injectors, traps, …

Supporting accelerator modeling:

- across DOE (HEP, BES, NP, FES, DNN) and beyond (CERN, DESY, KEK, …).

All codes share Particle-In-Cell loop at core
(as do other accelerator & plasma codes elsewhere)

Strategy relies then on small kernel

- NESAP work should benefit other BLAST codes, and beyond (OSIRIS, Synergia, …)

- Warp is a complex code:
  - ~150k lines FORTRAN + ~100k lines Python

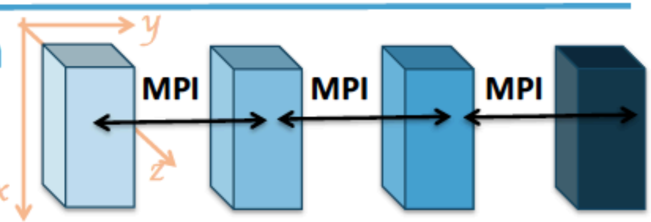
→ Particle-In-Cell Scalable Applications Resources (PICSAR)
  - initiated with kernel of Warp's Particle-In-Cell main loop (all FORTRAN)
    - optimization independently of Warp complexity and legacy
  - open source repository for collaborative development and distribution with other codes/groups
    - licensing is underway

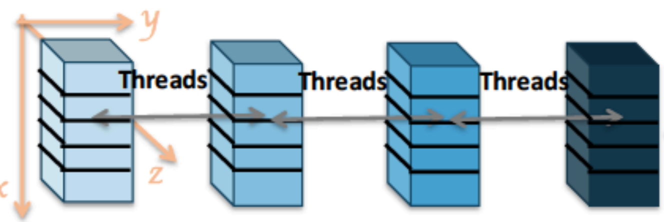
(derived class of PIC solver enables usage of new optimized routines from Warp)

Novel pre-exascale supercomputers require restructuration with "multi-level parallelism"

To run effectively on future systems

- **Manage Domain Parallelism**
  - independent program units; explicit
- **Increase Thread Parallelism**
  - independent execution units within the program; generally explicit
- **Exploit Data Parallelism**
  - Same operation on multiple elements
- **Improve data locality**
  - Cache blocking; Use on-package memory

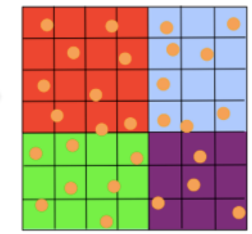




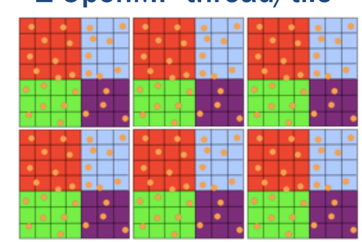

```
|--> DO I = 1, N
|        R(I) = B(I) + A(I)
|--> ENDDO
```

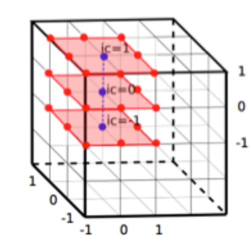
Particle-In-Cell

Domain decomposition
1 MPI task/domain



Subdomain tiling
1 OpenMP thread/tile





U.S. DEPARTMENT OF ENERGY | Office of Science

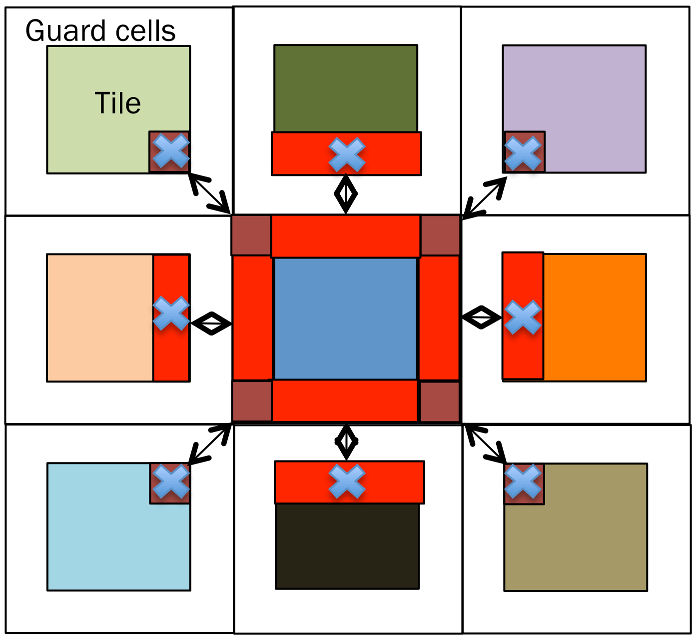
- 12 -

BERKELEY LAB

Courtesy Katherine Riley, FES Exascale Review, 2016

Initial OpenMP scheme gave poor scaling

One MPI SUBDOMAIN



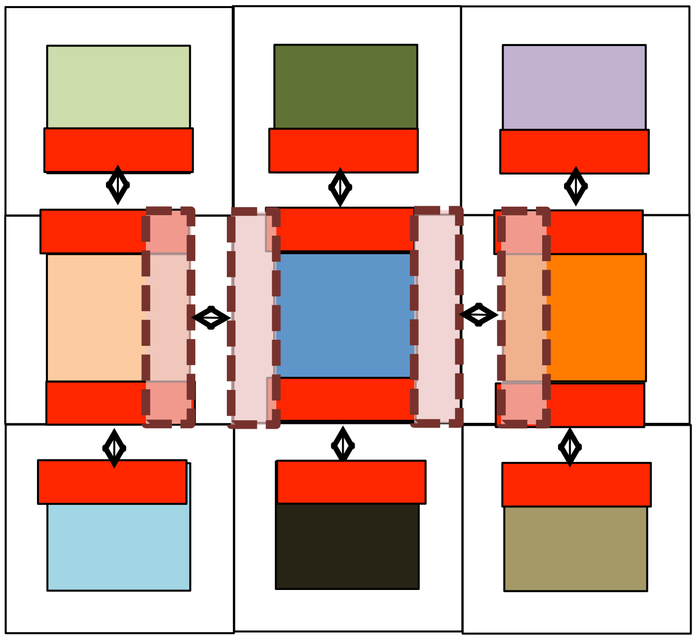
One big loop with one thread/tile

- Two adjacent tiles may write to the same location (conflicts)
- Adding $OMP REDUCTION clause not sufficient
- Poorly scales with # OpenMP threads:

  5x on Edison on one socket (12cores)

Novel OpenMP scheme to improves scaling

One MPI SUBDOMAIN



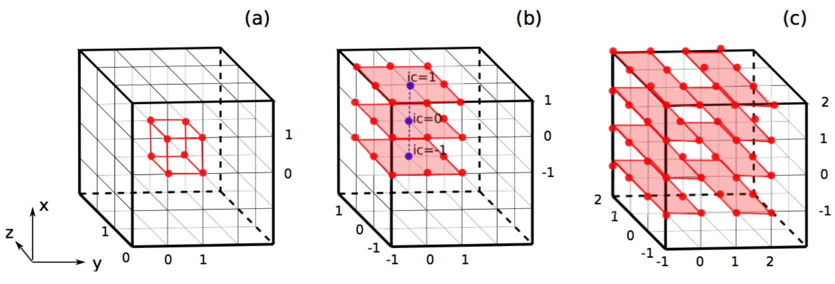
Main loop split in 4:

- Loop 1: Each threads write on its center part
- Loop 2: Each thread write in +/-X direction
- Loop 3: Each thread write in +/-Y direction
- Loop 4: Each thread write in +/-Z direction

OpenMP scaling: 5x -> 11x! (on Edison)

Novel vectorization algorithm leads to >2x speedup on charge/current deposition routines

- Previous algorithms developped on vector supercomputers in 70s-90s do not work
- Novel algorithm developped* (implemented in Warp/PICSAR)

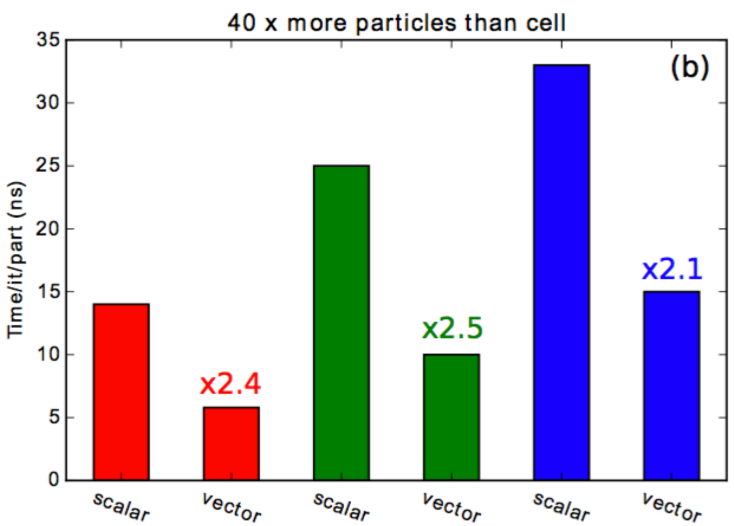


linear quadratic cubic

*H. Vincenti, R. Lehe, R. Sasanka, J-L. Vay, « An efficient and portable SIMD algorithm for charge/current deposition in Particle-In-Cell codes », arXiv:1601.02056, submitted to Comp. Phys. Comm.

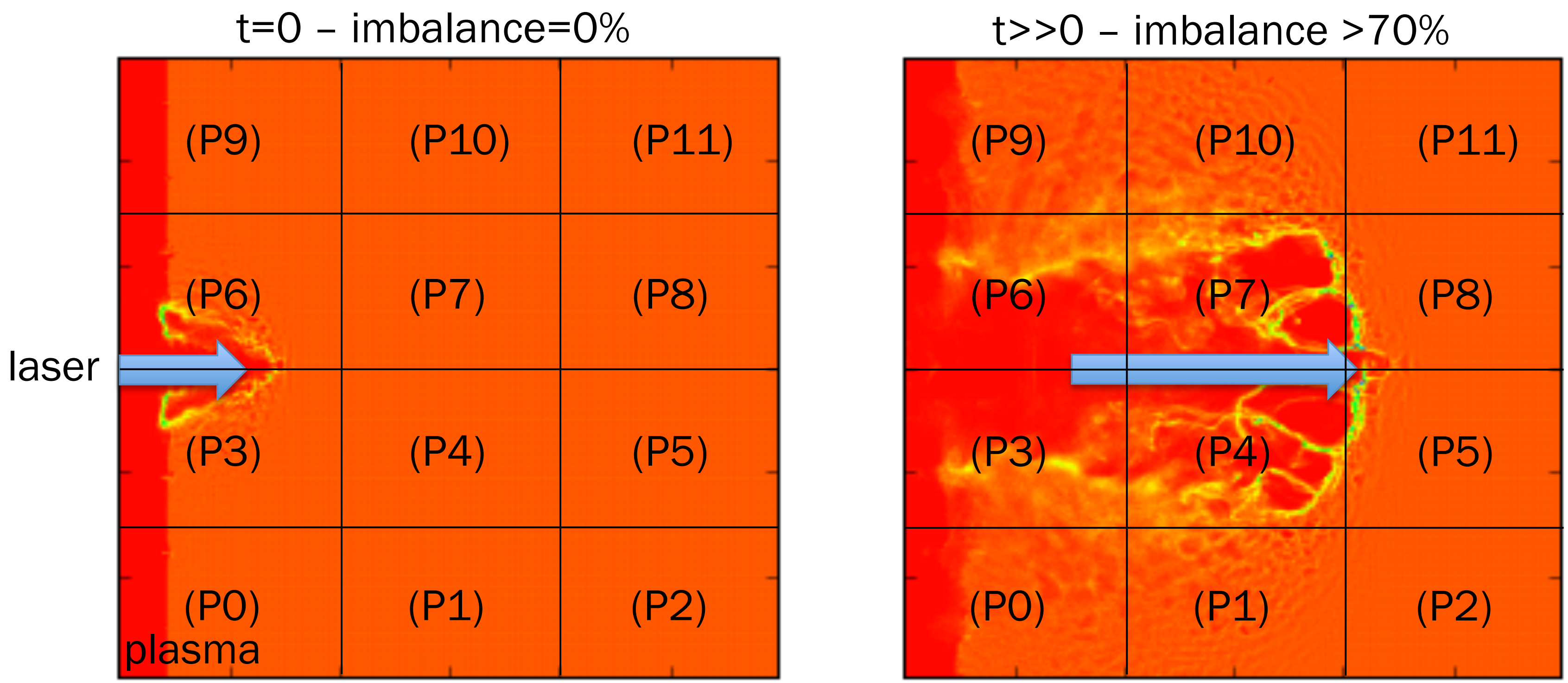
➤ Benchmarks on Cori (Haswell CPU)

Tests demonstrate speedups >2x (max. theoretical=4)



Load imbalance develops during simulations

- Distribution of particles, initially regular, becomes very irregular
- E.g.: laser-driven ion simulation with 12 cores

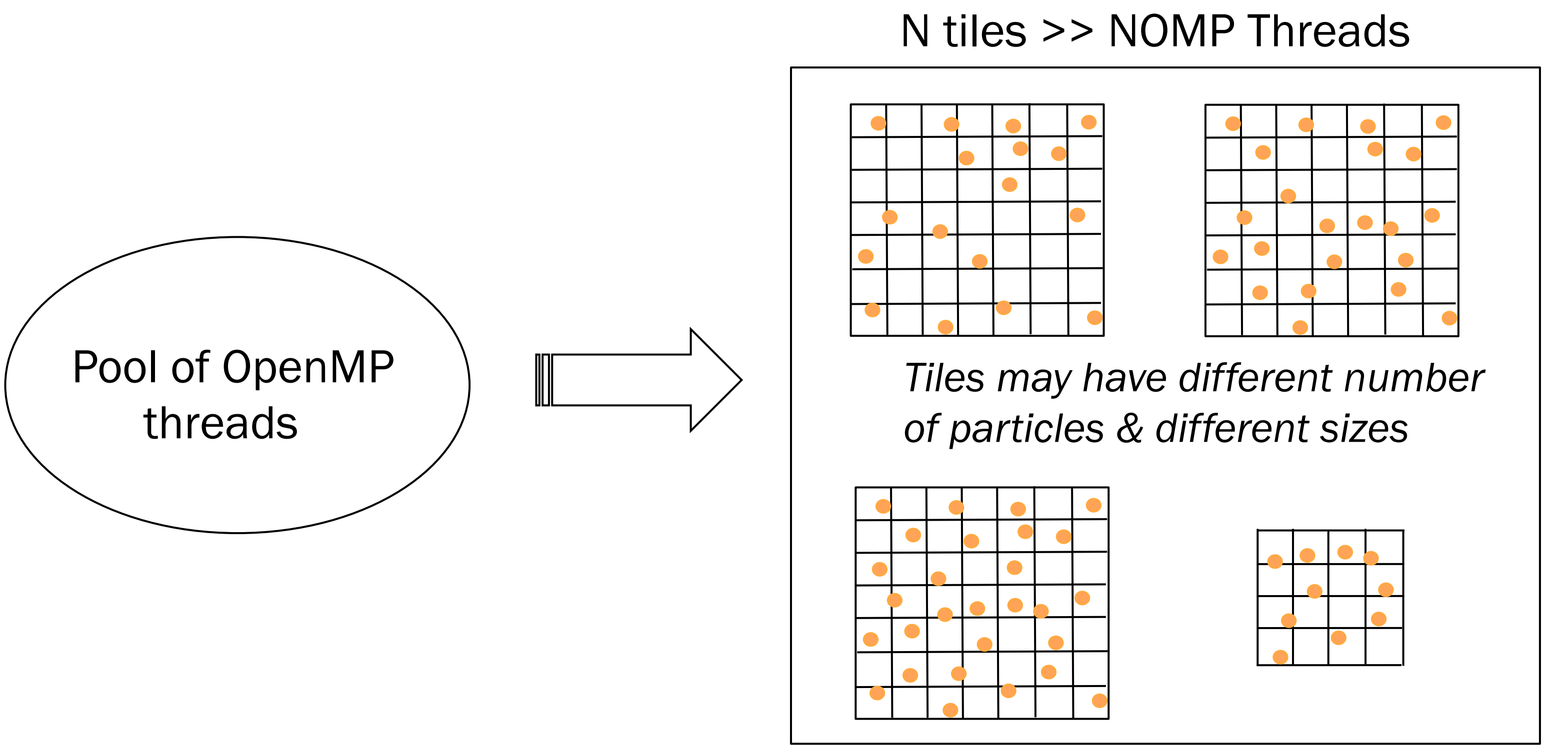


Imbalance (%) = 100x(MAXTIME(Pi)-MINTIME(Pi))/MINTIME(Pi)

Irregularity also affects balance between tiles

➔ needs for intra- & inter-node dynamic load balancing (DLB)

Intranode DLB by having N tiles >> NOMP threads



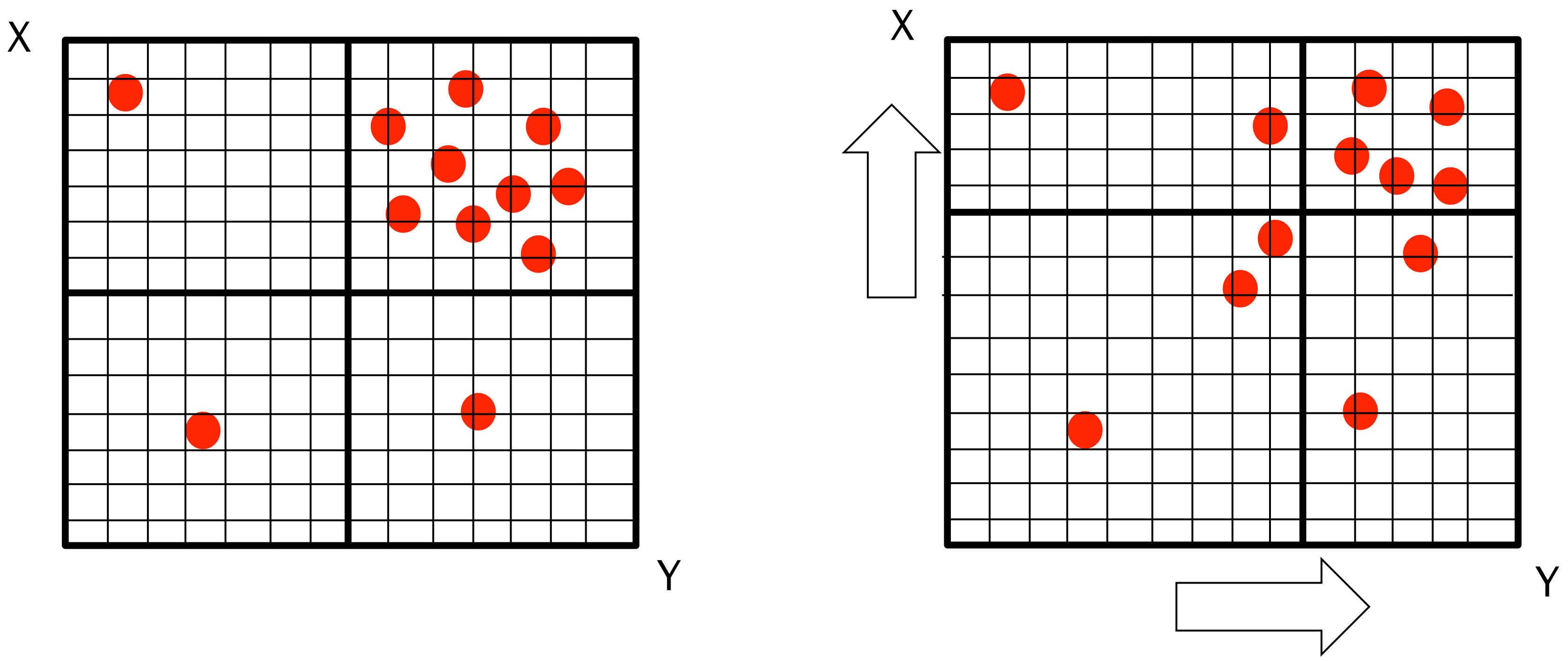
- DLB implemented using $OMP SCHEDULE(runtime) with runtime="guided" or "dynamic"

- Initial testing show better results with "dynamic"

Internode DLB by moving domain boundaries

Simple strategy:

(i) Project work load (field solvers+particle routines) along X,Y,Z
(ii) Compute new CPU boundaries along X,Y,Z
(iii) Exchange particles and fields between CPUs



Testing of efficiency on large simulations is underway

More in talk by A. Bhagatwala tomorrow

I/O, data analysis and visualization

Large scale HPC simulations need efficient I/O, analysis and visualization

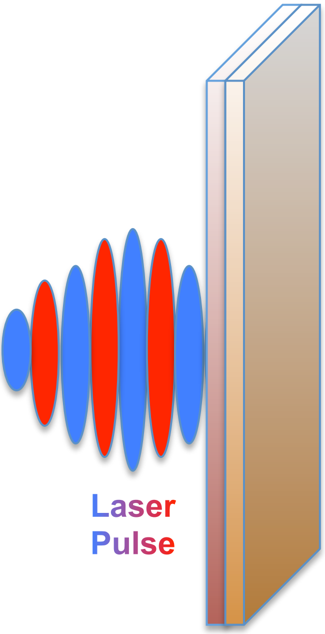
- I/O, analysis and visualization can be bottleneck
- collaborations on HDF5 and ADIOS parallel I/O
- collaboration on new data layout specifications for PIC
  - https://github.com/openPMD
  - http://www.openpmd.org
- collaboration on in situ data analysis and visualization
  - https://bitbucket.org/berkeleylab/warpiv

Presentation by B. Loring on Monday

In situ analysis & viz. complements raw data dumps

- Dumps of particles and fields are needed for post-processing
  - but limited by parallel I/O efficiency & amount of data generated
- *In situ analysis* enables increased frequency
  - down-samples, derived quantities, reduced geometry etc
  - focused analysis, write exactly what you (think you) need
- *In situ visualization* enables maximum temporal resolution and I/O reduction of multi-dimensional datasets
  - reduces problem sized data to image sized data
  - fast, small, serial I/O, doesn't impact file system.

Example: laser-driven ion beam acceleration

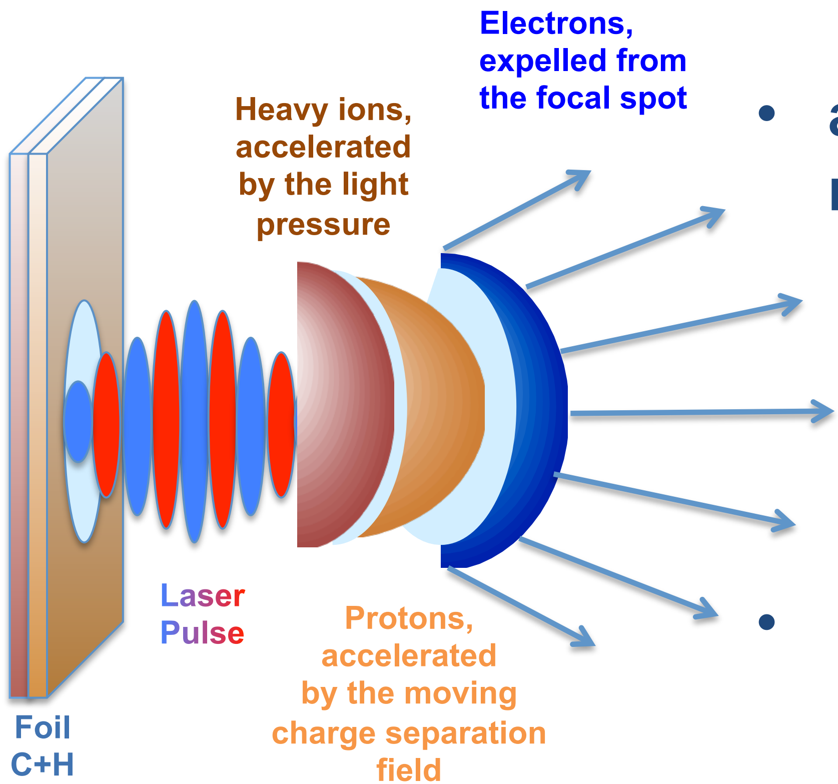


- active area of research with many applications!
  - including cancer therapy
    - precise control over depth at which energy is deposited
      ➔ less tissue damage
    - relatively compact and cheap
- HPC simulations are essential
  - used to develop and prove theory
  - when mature, will be used to model experimental apparatus

simple idea: use a laser to create a plasma. charge separation accelerates protons

S. S. Bulanov, et al. "***Accelerating monoenergetic protons from ultrathin foils by flat-top laser pulses in the directed-coulomb-explosion regime,***" *Phys. Rev. E, vol. 78, p. 026412, Aug 2008.*

Example: laser-driven ion beam acceleration

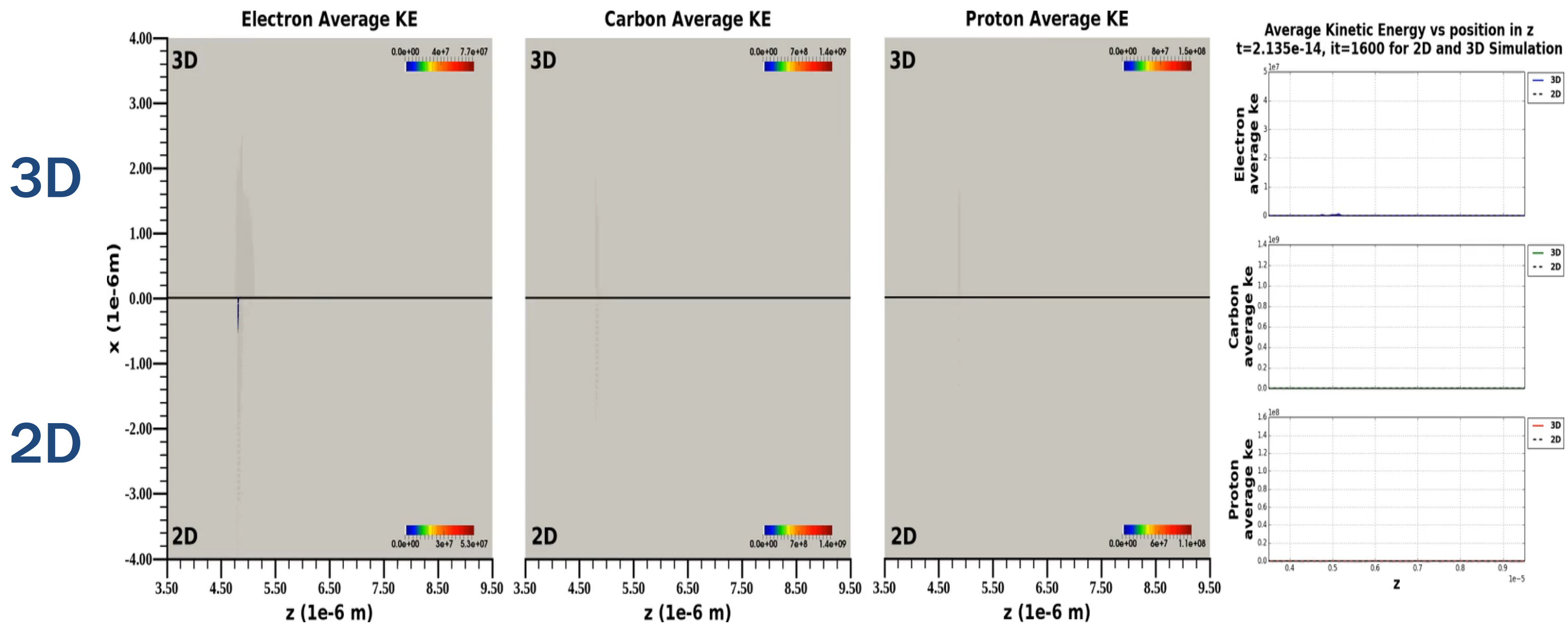


simple idea: use a laser to create a plasma. charge separation accelerates protons

- active area of research with many applications!
  - including cancer therapy
    - precise control over depth at which energy is deposited
      ➔ less tissue damage
    - relatively compact and cheap
- HPC simulations are essential
  - used to develop and prove theory
  - when mature, will be used to model experimental apparatus

*S. S. Bulanov, et al. "**Accelerating monoenergetic protons from ultrathin foils by flat-top laser pulses in the directed-coulomb-explosion regime,**" Phys. Rev. E, vol. 78, p. 026412, Aug 2008.*

3D simulations are needed to capture correct physics

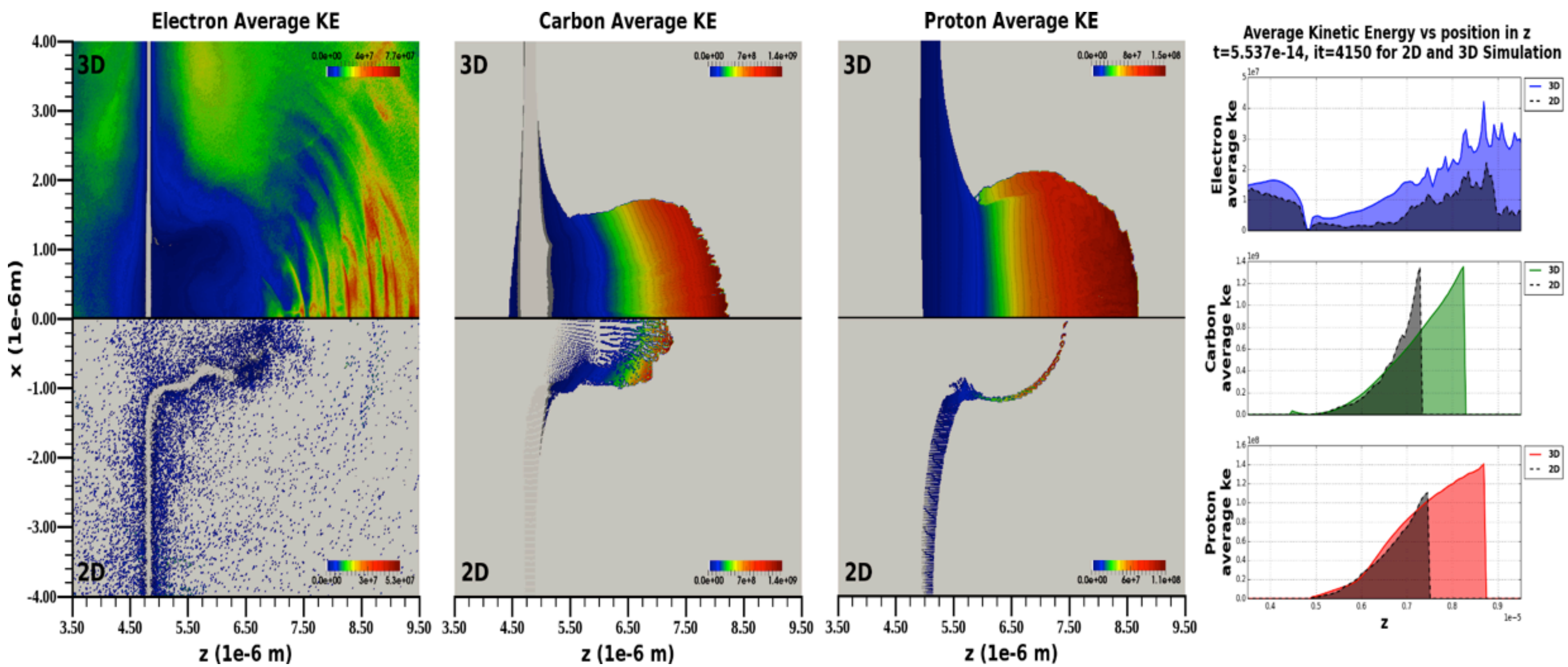


- 2D is qualitatively similar, but in 3D:
  - higher energies are reached
  - the beam propagates faster and further over the same time period.
- In-situ data analysis & viualization enables high frequency projections & histories of derived quantities

3D simulations are needed to capture correct physics

3D

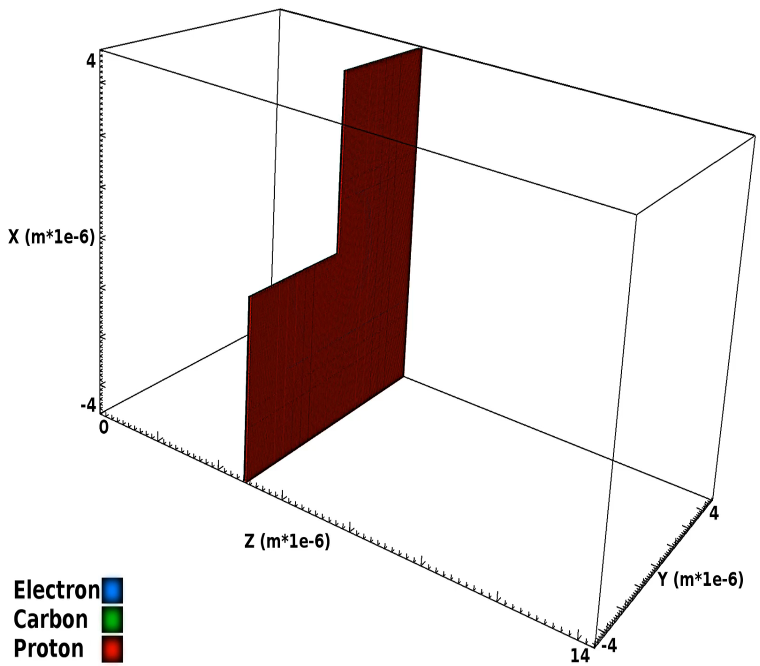
2D



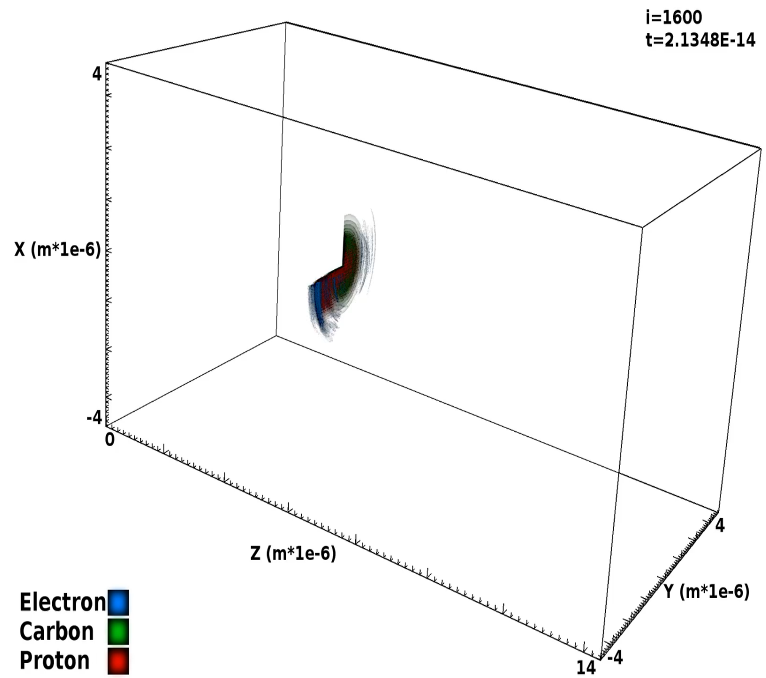
- 2D is qualitatively similar, but in 3D:
  - higher energies are reached
  - the beam propagates faster and further over the same time period.
- In-situ data analysis & viualization enables high frequency projections & histories of derived quantities

In situ enables high-frequency visualization of isosurfaces

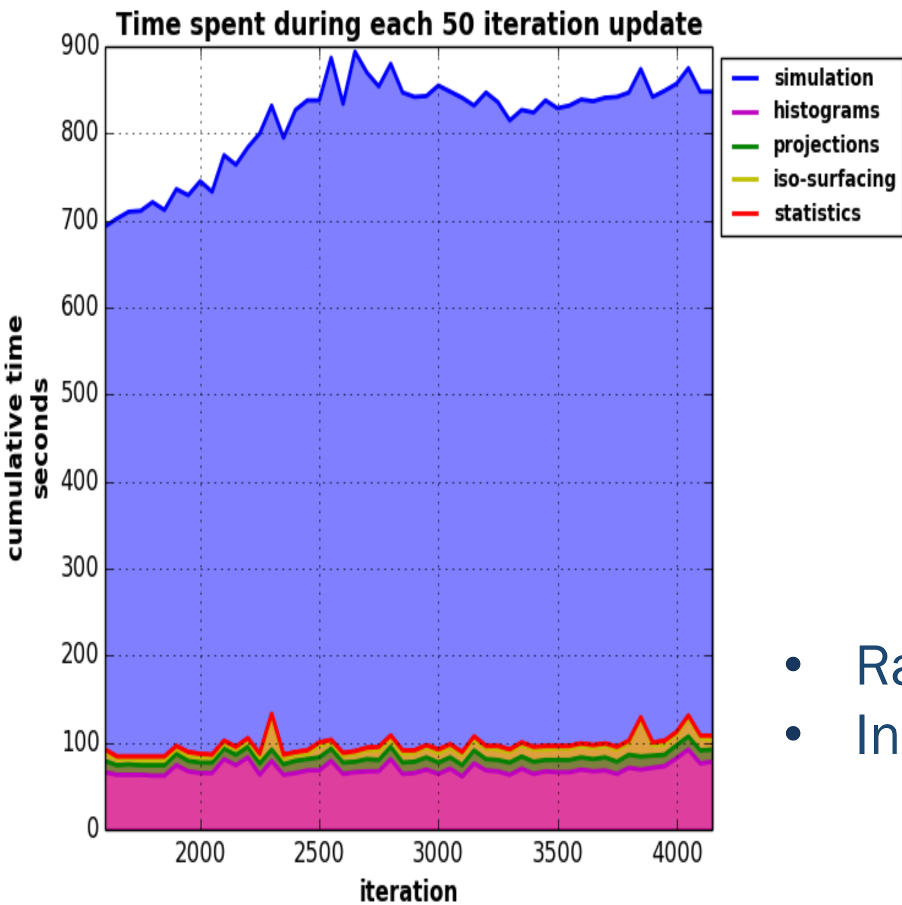
Particle Density



Kinetic Energy



In situ analysis and visualization: small portion of runtime and small I/O footprint



Category	Written to disk
histograms	4.5 MB
projections	796. MB
iso-surfaces	40.8 MB
statistics	3.5 MB
total	**841. MB**

- Raw data **for** same analysis ~ **3.4 TB.**
- In situ ➔ **4033×** reduction in I/O.

Summary

- Particle accelerators are essential tools of science, medicine, industry and security
- Plasma-based methods on the brink to deliver much smaller, cheaper accelerators, with profound impacts
- HPC already essential and can play key role in the development of new technologies
- Ultimate goal is real-time virtual prototyping of entire accelerators
- Preparation toward exascale is vital and underway (NESAP)
- Novel in situ analysis and visualization tools will complement standard parallel I/O for maximum utilization

NERSC is an essential partner of our research program and we look forward to our continued partnership and future co-discoveries!