Finite Difference Time Domain (FDTD) Electromagnetic Modeling Efforts of the ComPASS SciDAC (and related SBIR) Projects

J.R. Cary,[1,2] K. Amyx,[1] T. Austin,[1] *D. Bruhwiler,*[1]
B. Cowan,[1] P. Messmer,[1] P. Mullowney,[1] K. Paul,[1]
P. Stoltz,[1] D. Smithe,[1] S. Veitzer[1] & G. Werner[1,2]

1. Tech-X Corp.
2. University of Colorado

Large Scale Computing and Storage Requirements for High Energy Physics (An HEP/ASCR/NERSC Workshop)

Rockville, MD, November 12, 2009

Supported by the US DOE Office of Science, Office of High Energy Physics, via the SBIR & SciDAC programs; and also by ASCR, Nuclear Physics & Basic Energy Sciences via the SBIR program; used NERSC supercomputing center.

Motivating Science & Applications

- Supporting multiple NERSC projects:
  - Community Petascale Project for Accelerator Science and Simulation
  - Particle simulation of laser wakefield particle acceleration
  - Simulation of photonic crystal structures for laser driven particle acceleration

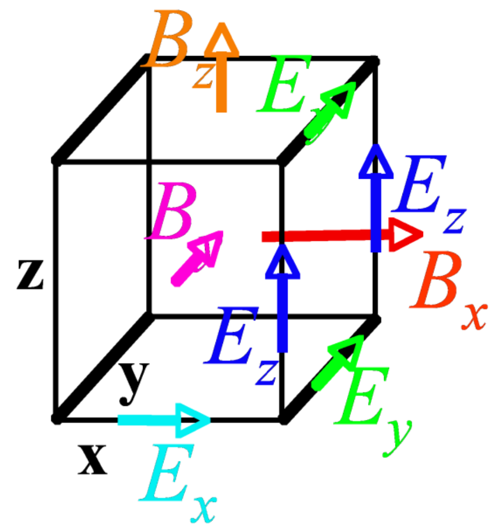
- Scientific objectives for the next 3-5 years: rapid high-fidelity simulation & design of a wide range of accelerating structures
  - Superconducting radio-frequency (SRF) accelerating cavities
    - calculation of frequencies, Q's, modes (fundamental & high-order); surface heating
    - multipacting – move from analysis to designs that mitigate the problem
    - DOE/HEP applications include LHC, Project X, ILC
  - Normal conducting (warm) RF cavities and waveguides
    - breakdown – move from analysis to designs that mitigate the problem
    - simulate "magnetic insulation" of novel RF cavities for muon acceleration
    - DOE/HEP applications include muon collider, RF power transport, CLIC-like concepts
  - Dielectric structures (advanced concepts)
    - high-gradient, laser-driven photonic band gap (PBG) accelerating cavities
    - novel, larger-scale RF structures with ultra-high Q, ultra-low wakefields
  - Multi-physics capabilities are required, especially surface physics
    - coupling electromagnetics to surface heating & heat transport
    - electron-wall interactions

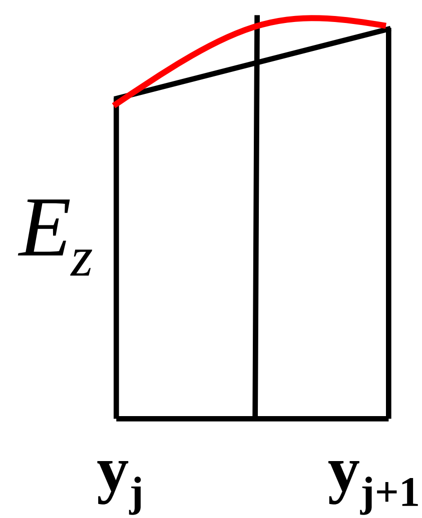
TECH-X CORPORATION

FDTD computations are based on the simple, fast Yee algorithm

- No matrix inversions
- Manifestly stable
  - Symmetric update matrix
- Works well with particles (PIC)
  - <u>The choice</u> of PIC codes
- Parallelizes well
  - Only boundary information exchanged between domains

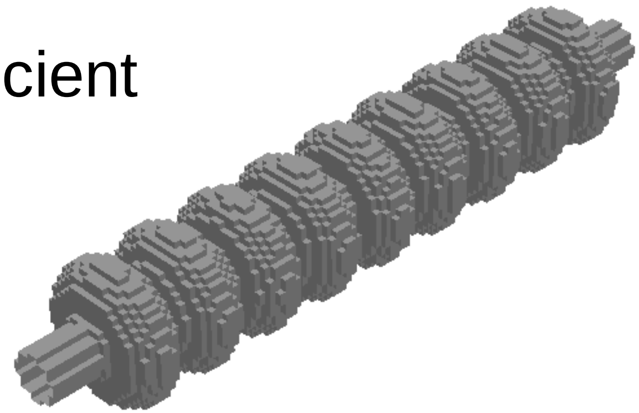
– Higher-order versions exist

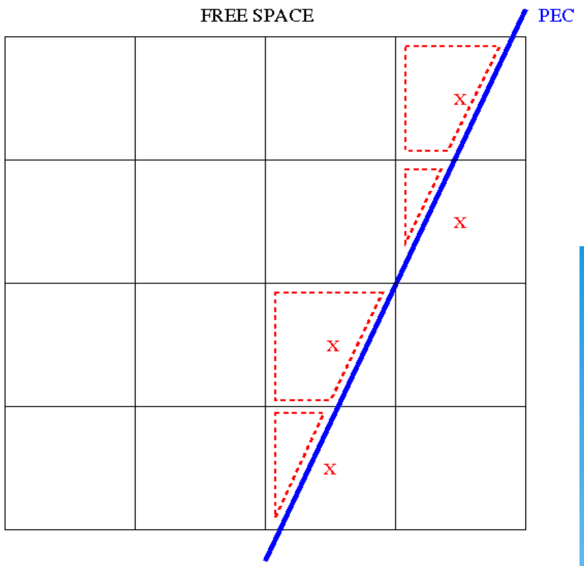


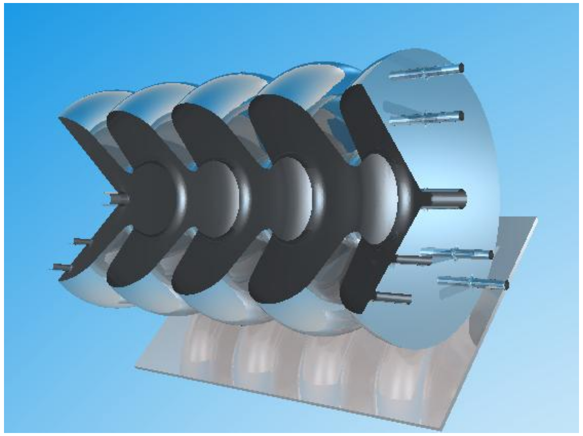


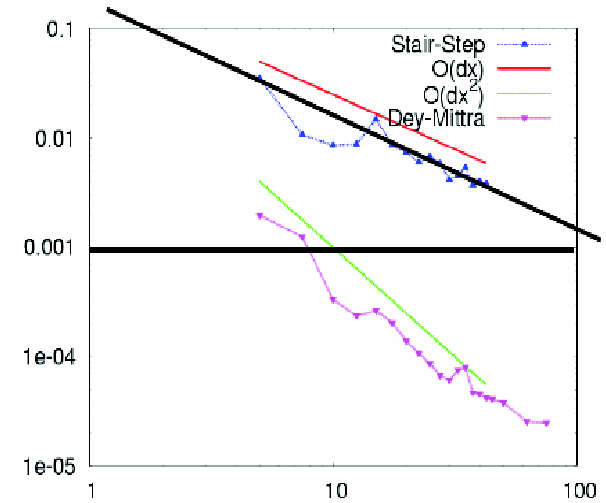
Curved structures accurately modeled with embedded boundaries

- Stairstep was not accurate
- Dey-Mittra found to give sufficient accuracy









Improved algorithms are being developed

- *Broadly filtered diagonalization method:*

IER

Journal of Computational Physics 227 (2008) 5200–5214

www.elsevier.con

Extracting degenerate modes and frequencies from time-domain simulations with filter-diagonalization [q]

Gregory R. Werner [a,*], John R. Cary [a,b]

[a] Center for Integrated Plasma Studies, University of Colorado, Boulder, CO 80309, United States
[b] Tech-X Corporation, Boulder, CO 80303, United States

Received 22 May 2007; received in revised form 20 December 2007; accepted 27 January 2008
Available online 5 February 2008

- *Symmetric dielectric update algorithms:*

Journal of Computational Physics 226 (2007) 1085–1101

www.elsevier.(

A stable FDTD algorithm for non-diagonal, anisotropic dielectrics [q]

Gregory R. Werner [a,*], John R. Cary [a,b]

[a] Center for Integrated Plasma Studies, University of Colorado, Boulder, CO 80309, United States
[b] Tech-X Corporation, Boulder, CO 80303, United States

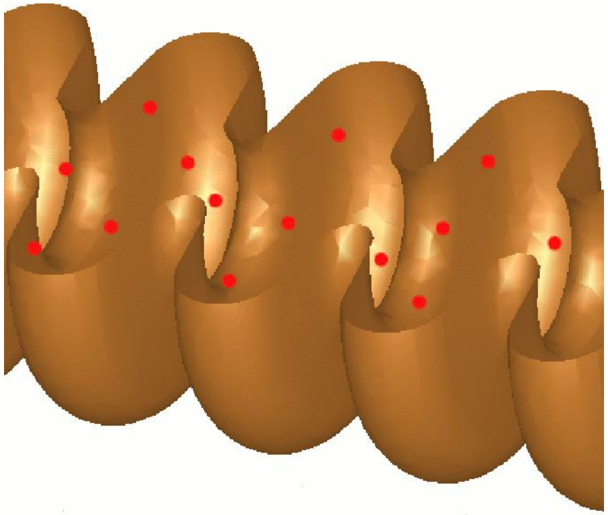
Received 9 February 2007; received in revised form 4 May 2007; accepted 7 May 2007

VORPAL

TECH-X CORPORATION

Broadly filtered diagonalization: time-domain codes become frequency domain

- Traditional method to obtain frequencies from time-domain codes:
  - Excite one mode with narrow band
  - Measure FFT peak or zero crossing
  - Cannot distinguish degeneracies

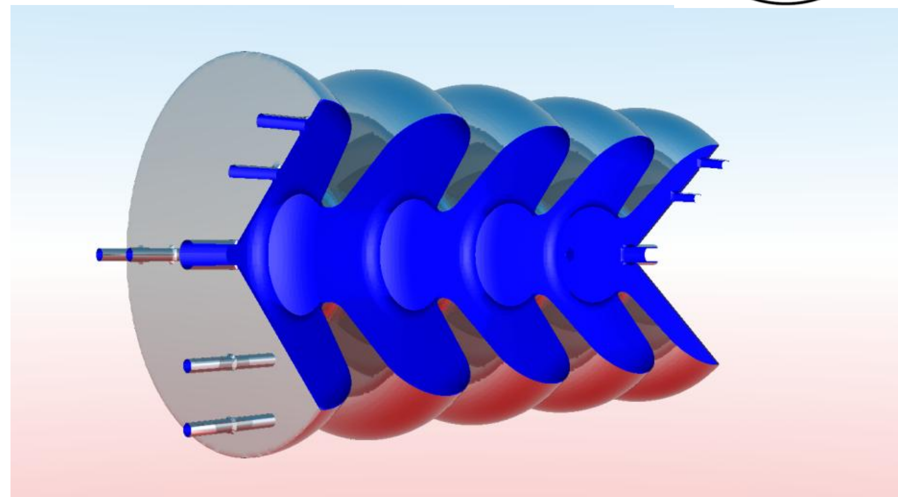


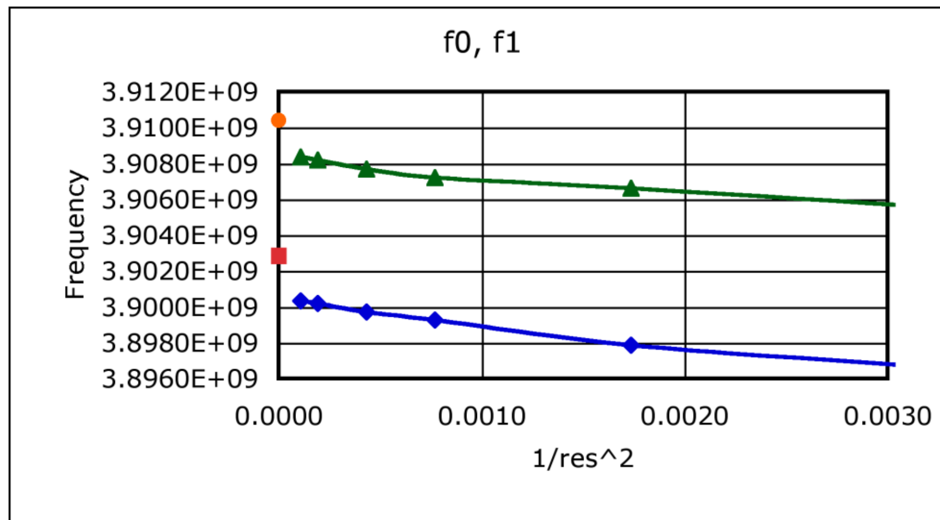
- Broad filtering
  - Excite collection of modes in a frequency band
  - Collect data on a subspace
  - One application of operator gives small relative eigenvalue problem
  - Singular value decomposition determines the linearly independent subspace
  - Degeneracies found

Eliminates requirements for retention of multiple eigenvectors for eigenvalue solving

Validation with Experimental SRF Cavity Measurements have been Successful

- Collaboration with Fermilab
- A15 cavity for Kaon separator
  - Previous computations:
    - gave frequencies low by 5 MHz out of 4 GHz.
  - VORPAL computations (improved algorithm, parallel):
    - low by 2 MHz
    - verified against exact solutions
- Many attempts to understand discrepancy:
  - Model no holes? one? all?
  - Correct for dielectric of air?





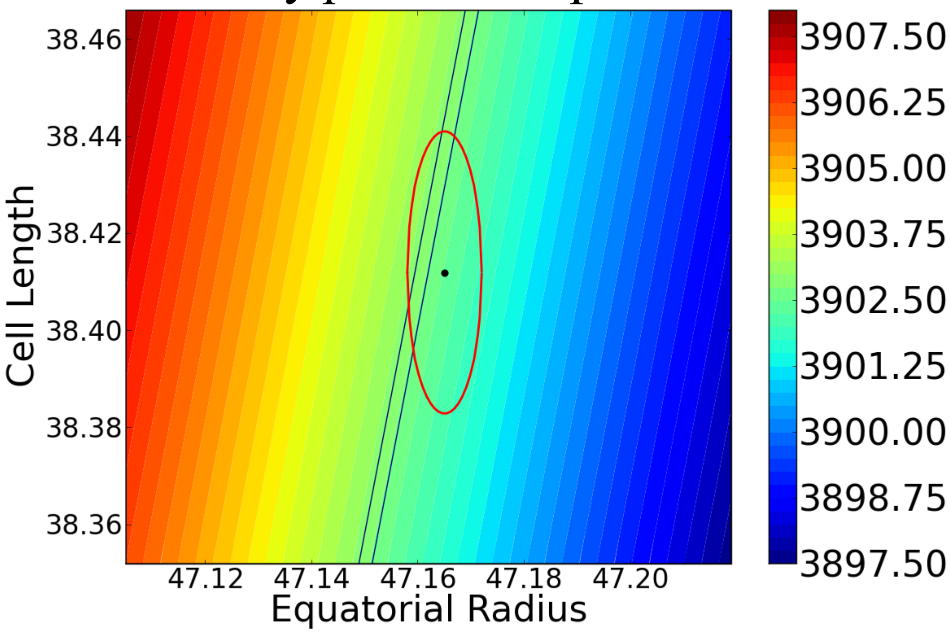
VORPAL

TECH-X CORPORATION

A15 cavity validation study with Fermilab identified error in previous measurements

- Reduce the equator radius by 0.001 inch
- Get agreement
- Ask designers to measure their cavities
- CORDEX (+ calipers) show error in cavity dimensions
- Corrected model agrees well with computation.

Frequency contours in cavity parameter space



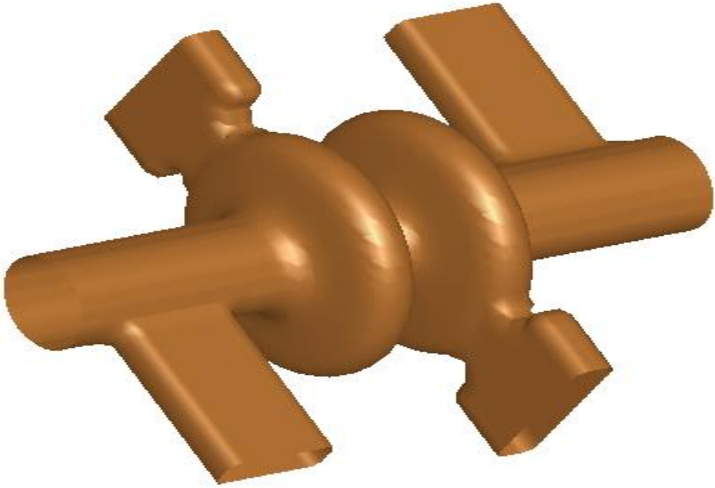
Overlap of dimensional error ellipse with computational and experimental frequency uncertainty shows validation

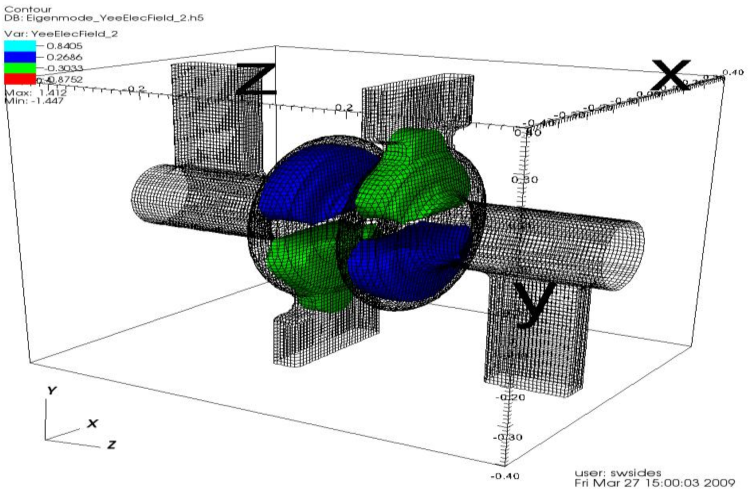
VORPAL

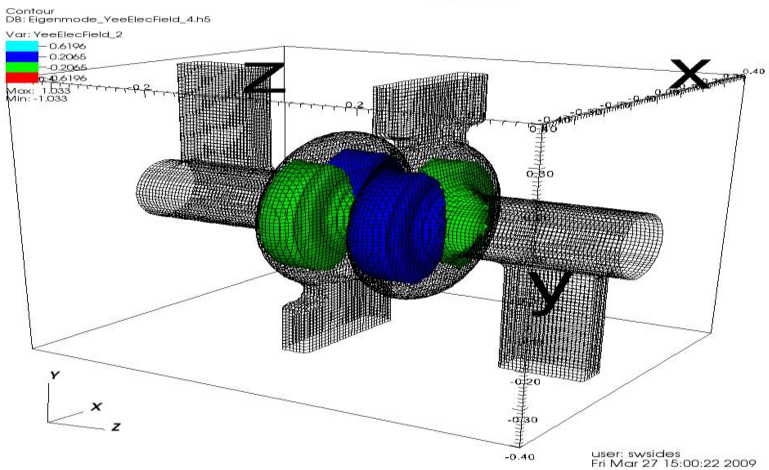
TECH-X CORPORATION

Multipactoring determines field limits in crab cavities

- LHC upgrade: introduce crab cavity to improve luminosity
- Jlab/Cockcroft splitting from waveguide replacement







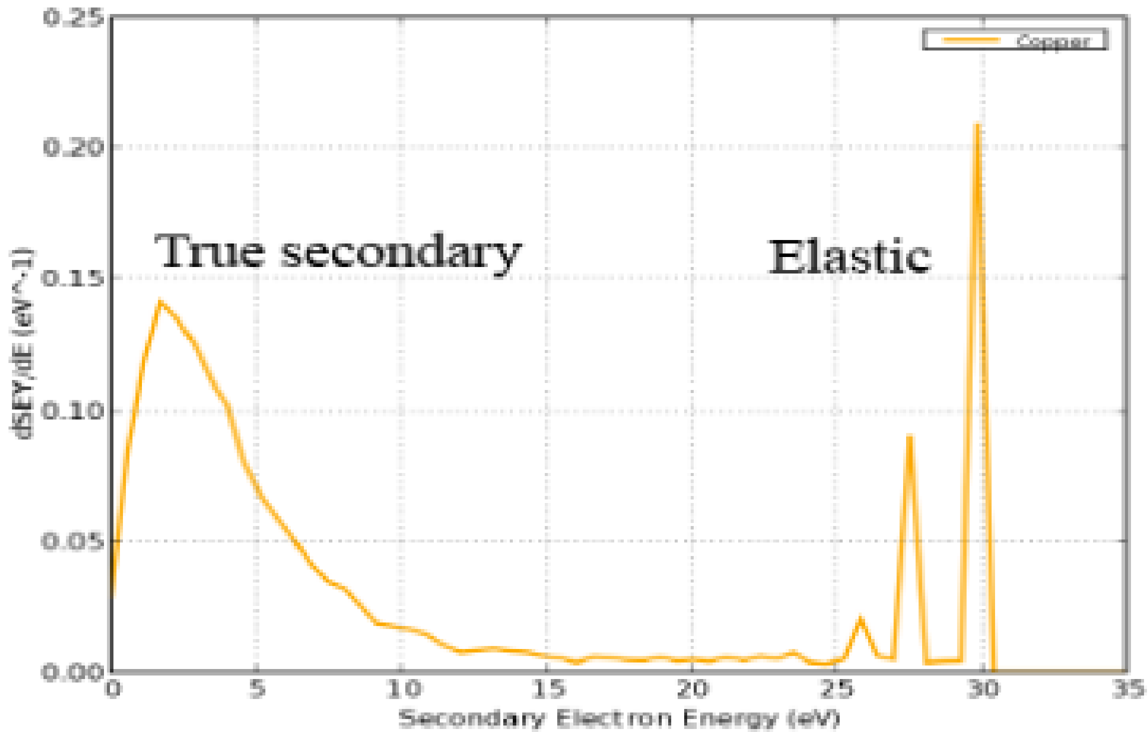
G. Burt, J. Smith (Cockcroft Institute)
H. Wang, K. Tian, R. Rimmer (Jefferson Lab)

VORPAL

TECH-X CORPORATION

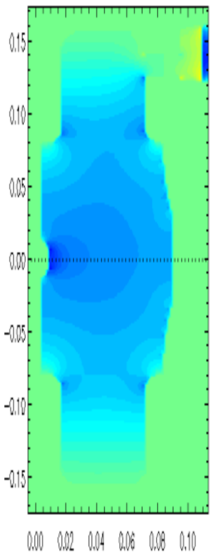
VORPAL's secondary electron emission models allow realistic multipacting studies

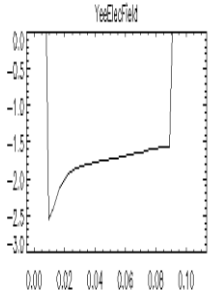
- Simple secondary emission – one secondary emitted at normal incidence
- Phenomenological model – true, diffuse and elastic secondaries
- M. A. Furman and M. Pivi, Phys. Rev. ST Accel. Beams 5, 124404 (2002)

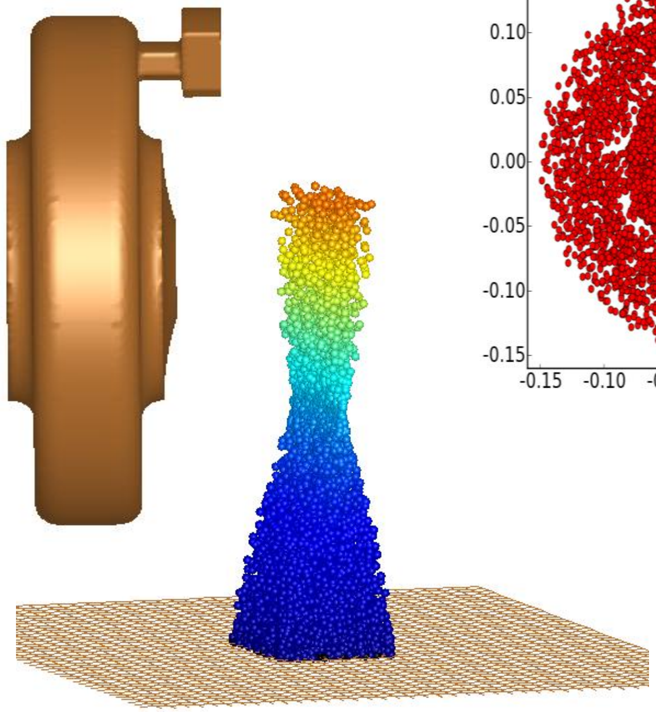


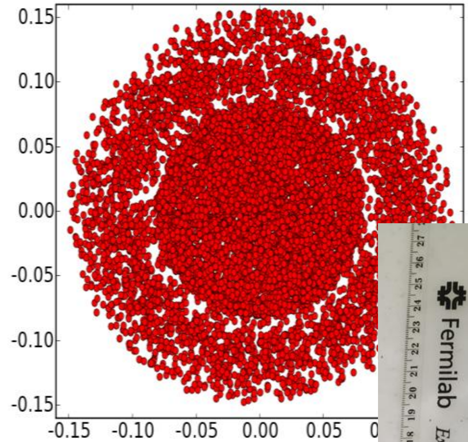
Available as 'txphysics' at http://www.txcorp.com/products/TxPhysics/
Library funded under SBIR. Made available in SciDAC codes.

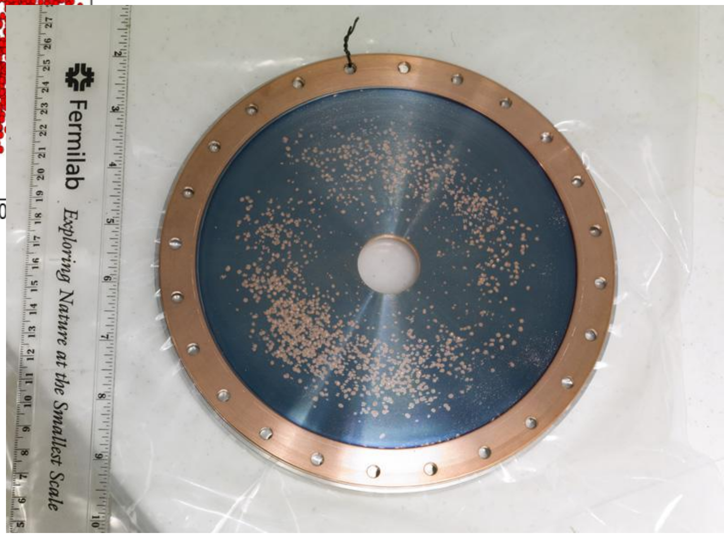
Multipacting analysis capability used to evaluate high-gradient muon collider cavities







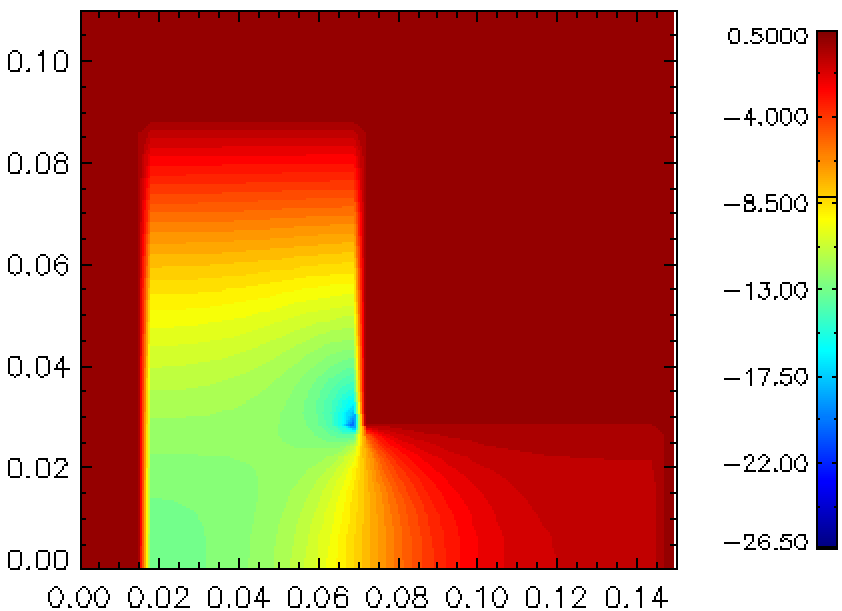




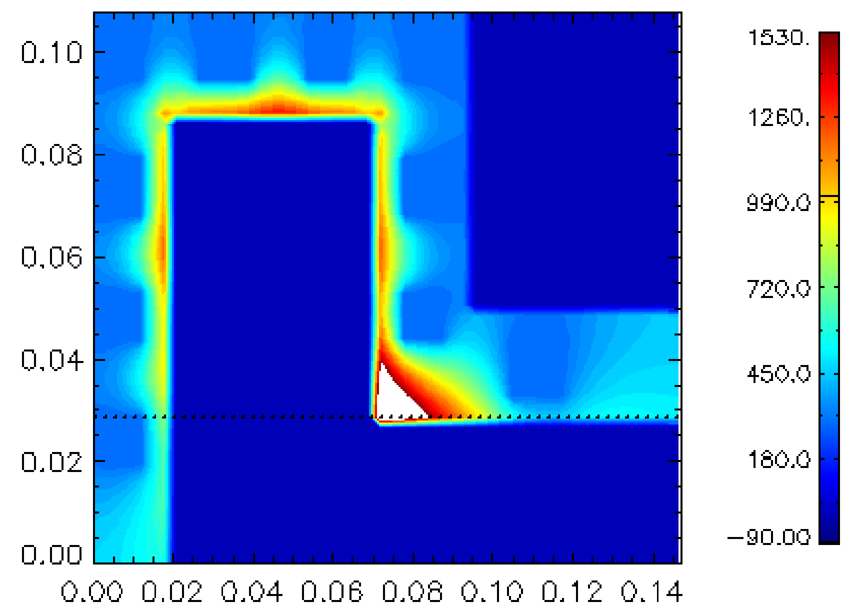
- VORPAL has accurate field models and physics-based surface emission algorithms (right)
- Multipacting and breakdown often coincident. Simulation shows multipacting where experiment shows breakdown.

Developing new multi-physics capabilities: combined EM & heat flow

- Simplified RF photocathode 'electron gun', with 10 coolant channels.
- EM and thermal updates proceed simultaneously in VORPAL
- Ohmic wall losses are communicated between EM and Thermal.
- The thermal timescale is artificially reduced to match the EM time-scale.
- R_s, k, and C are temperature dependant, e.g., non-linear.



EM: E_z Electric Field



Thermal: Temperature

VORPAL

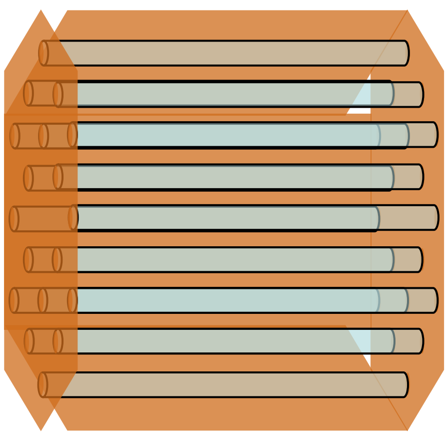
TECH-X CORPORATION

VORPAL is successfully modeling dielectric accelerating structures

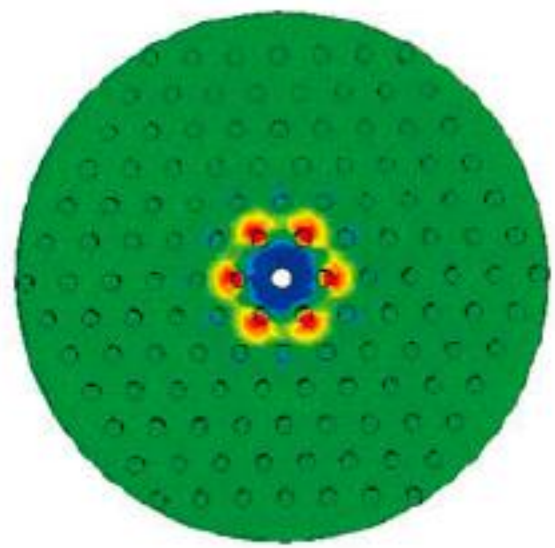
STUDY OF HYBRID PHOTONIC BAND GAP RESONATORS FOR PARTICLE ACCELERATORS

M. R. Masullo,[1] A. Andreone,[2] E. Di Gennaro,[2] S. Albanese,[3] F. Francomacaro,[3] M. Panniello,[3] V. G. Vaccaro,[3] and G. Lamura[4]

2486 MICROWAVE AND OPTICAL TECHNOLOGY LETTERS / Vol. 48, No. 12, December 2006 DOI 10.1002/mop

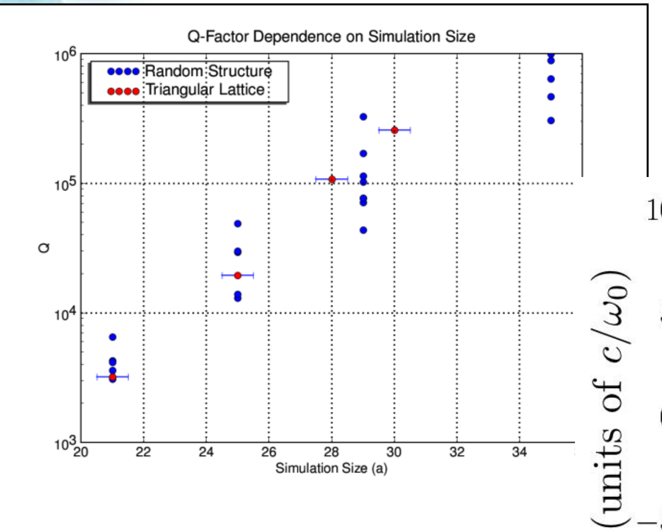


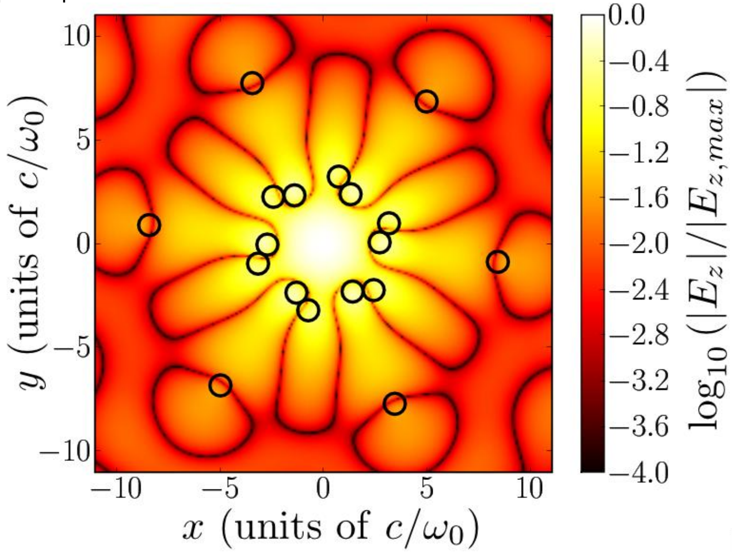
room temperature confirm the monomodal behavior, but the Q value is lower than expected (roughly 10^3). This is mainly due to



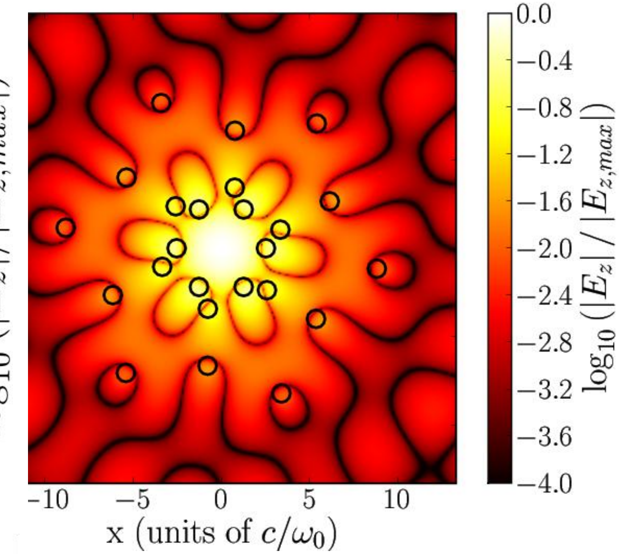
- 5 layers of (147) dielectric cylinders, yet Q ~ 10^3

Our optimizations found asymmetric systems with many fewer rods, yet larger Q (quality factor)





Final: Q ~ 11,000



Final: Q ~ 100,000

- Computations showed that symmetry breaking critical

Current HPC Requirements

- Architectures: ATLAS (LLNL), Franklin (NERSC),
  NY Blue (BNL), Intrepid (ALCF), Jaguar (ORLCF)
- Compute/memory load for typical production runs
  - 1-10 million cells; < 100,000 time steps
  - > $2x10^6$ cells/s; ~ 1,000 cores; < 100 MB/core
  - Efficient use of 16,000 cores shown; ~1,000 cells per core
- Data read/written
  - 3 to 9 doubles per cell → < GB files; ~50 dumps per run
  - particle data is highly variable
- Necessary software, services or infrastructure
  - parallel I/O support for HDF5
  - 3D viz (i.e. Visit server on Franklin)
- Current primary codes and their methods or algorithms
  - VORPAL, Cartesian mesh, FDTD
- Known limitations/obstacles/bottlenecks
  - I/O scaling

Broad usage at NERSC & LCFs

- VORPAL openly available to DOE collaborators at NERSC (Franklin), ALCF, ORLCF (Jaguar)
- Large number of users
  - 30 at NERSC under 5 different projects (repos)
  - 5-10 at ALCF under 2 different projects
- Large number of hours
  - NERSC: 6M hrs 2008, 2M by mid March for 2009
  - ALCF: 15M hrs 2008, 5M by mid March for 2009
- High concurrency (routine use at high processor counts)
  - NERSC, 2009: 8656 for average job
  - ALCF, 2009: 8192 cores typical
- Thanks to Katie Antypas of NERSC for data

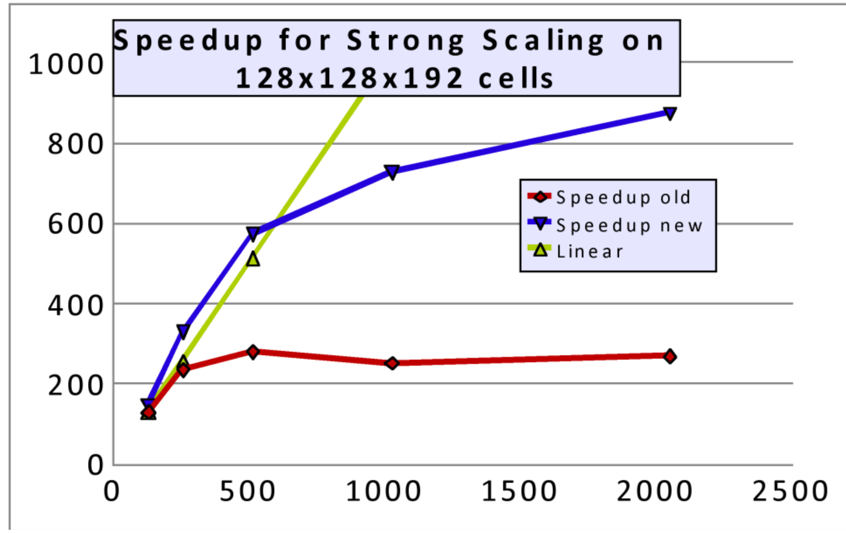
HPC Usage & Methods; next 3-5 Years

- Upcoming changes to codes/methods/approaches
  - implicit solvers; higher-order solvers (implicit and/or explicit);
  - coupling solvers for multiphysics (heat in wall, EM in vacuum);
  - improved on-the-fly data reduction to reduce/control I/O
  - single-core performance improvements; SIMD, single-precision
  - multiple grids; sophisticated static domain decomposition
- Changes to compute/memory load
  - 10x to 100x increase in mesh – up to 10^9 cells
  - for higher-order, implicit – increase in memory, not in mesh
- Changes to Data read/written
  - always more space is desired; working to reduce/slow growth
- Changes to necessary software, services or infrastructure
  - database support for parameter scans & design studies
- Anticipated limitations/obstacles/bottlenecks on 1000K PE system
  - file I/O (e.g. rapid dump/restore); fault tolerance; easy/rapid viz
- Strategy for dealing with multi-core/many-core architectures
  - actively porting to GPU

TECH-X CORPORATION

Developing: improved messaging for parallel performance enhancement

- Improved messaging: send only what is needed for FDTD
- Allows use of domains with only 3000 cells (before, 64000 cells)
- Consequences:
  - Time to solution increases by 20x if resources are available
  - Smaller problems can be addressed with high-performance computation



Break in strong scaling at 1000 procs or 3000 cell domains

– 120M cells can take advantage of 40k procs

Peter Messmer[1], Ben Cowan[1], George Bell[1], Keegan Amyx[1], Boyana Norris[2], John R. Cary[1]

[1]*Tech-X Corp.*, [2]*Argonne National Lab.*

Work supported by DOE ASCR SBIR Phase II DE-FG02-07ER84731 and by VORPAL customers

VORPAL

TECH-X CORPORATION

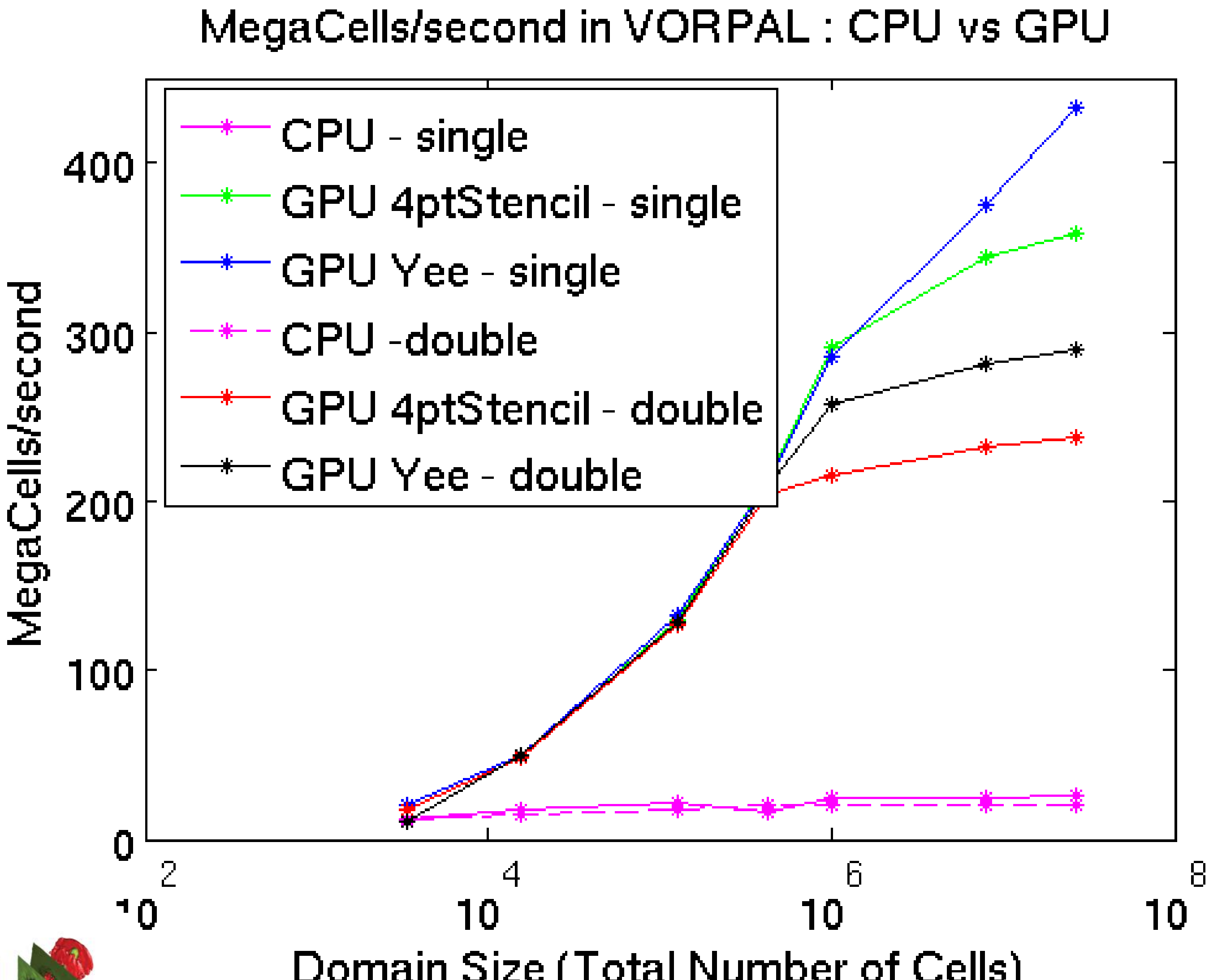
Recent GPU Developments

- VORPAL FDTD Simulations can be run across multiple GPUs. Speedup observed for large domains.
- Efficient data structures will improve messaging between CPU↔GPU → better performance across multiple devices for smaller domains

Future Work and Projections

- Dey-Mittra cut-cell algorithms will be implemented soon
  → highly resolved and efficient RF cavity simulations on a GPU.
- Move particle push onto GPU.
- Multi-GPU machines will
  - greatly accelerate current simulations or
  - enable more highly resolved simulations with greater fidelity
- GPUs are an energy-efficient, cost-effective tool for rapid plasma and EM simulations.

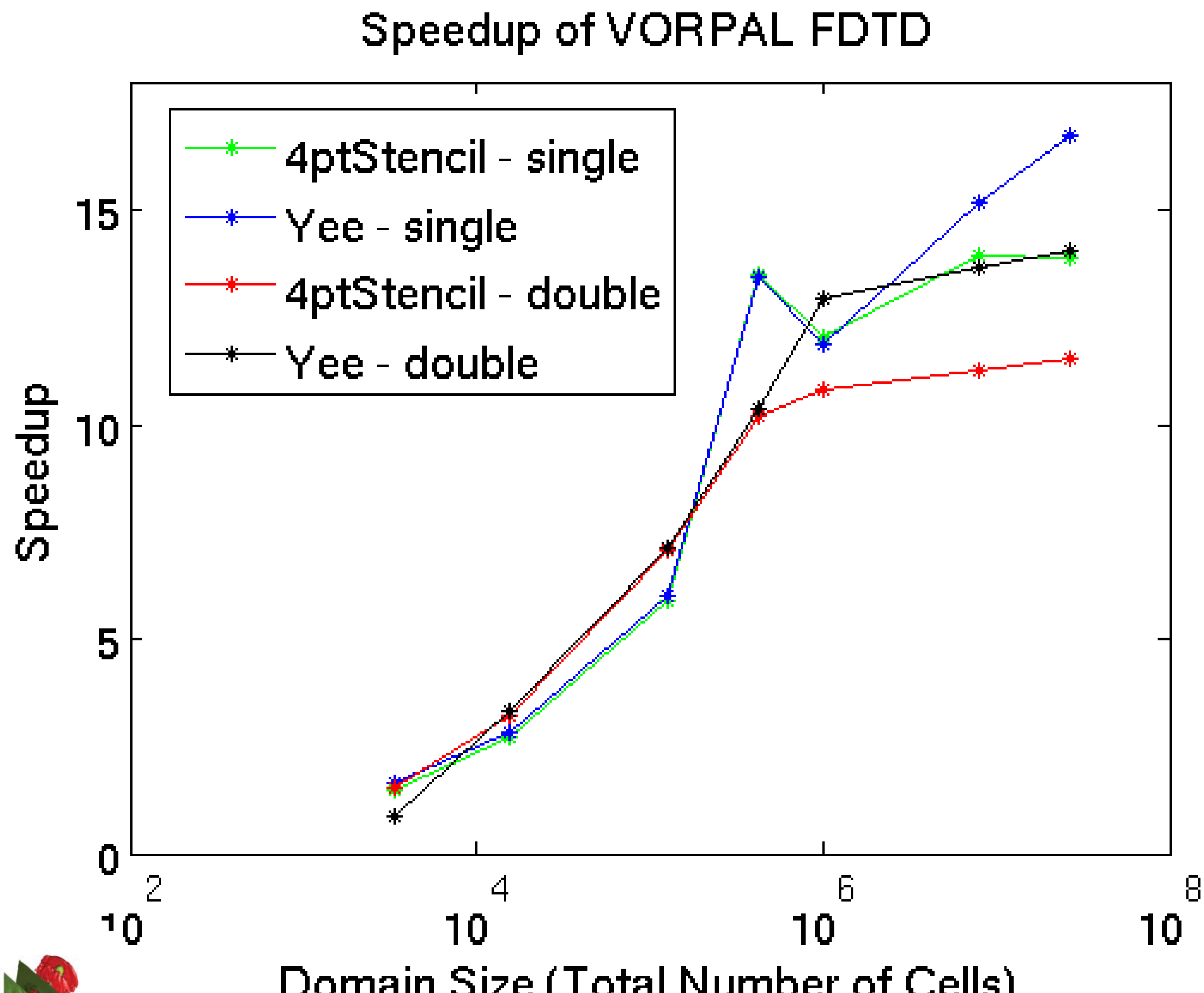
Timings: FDTD Performance on a single GPU or CPU



VORPAL

TECH-X CORPORATION

Speedup of GPU kernels : 1 GPU vs 1 CPU



Summary

- Recommendations on NERSC architecture, system configuration and associated service requirements needed for your science:
  - database support for parameter scans & design studies
  - support task-based parallelism, so that many large, parallel jobs can be executed (w/ optimizer?) for design of accelerators
  - primary viz support should target largest system (Hopper)
- What significant scientific progress could you achieve over the next 5 years with access to ~50X NERSC resources?
  - move from large, single run work flow to real design activities, including error analysis, to reduce cost & risk for future facilities
  - multi-physics EM modeling could yield smaller/cheaper accelerators:
    understand multipactoring → high-gradient SRF cavities
    understand RF breakdown → higher RF power availability
  - dielectric structure simulations could enable the development of fundamentally new accelerator hardware with much higher gradients