Parallel Mesh Partitioning at SLAC

Michael M. Wolf

Stanford Linear Accelerator Center

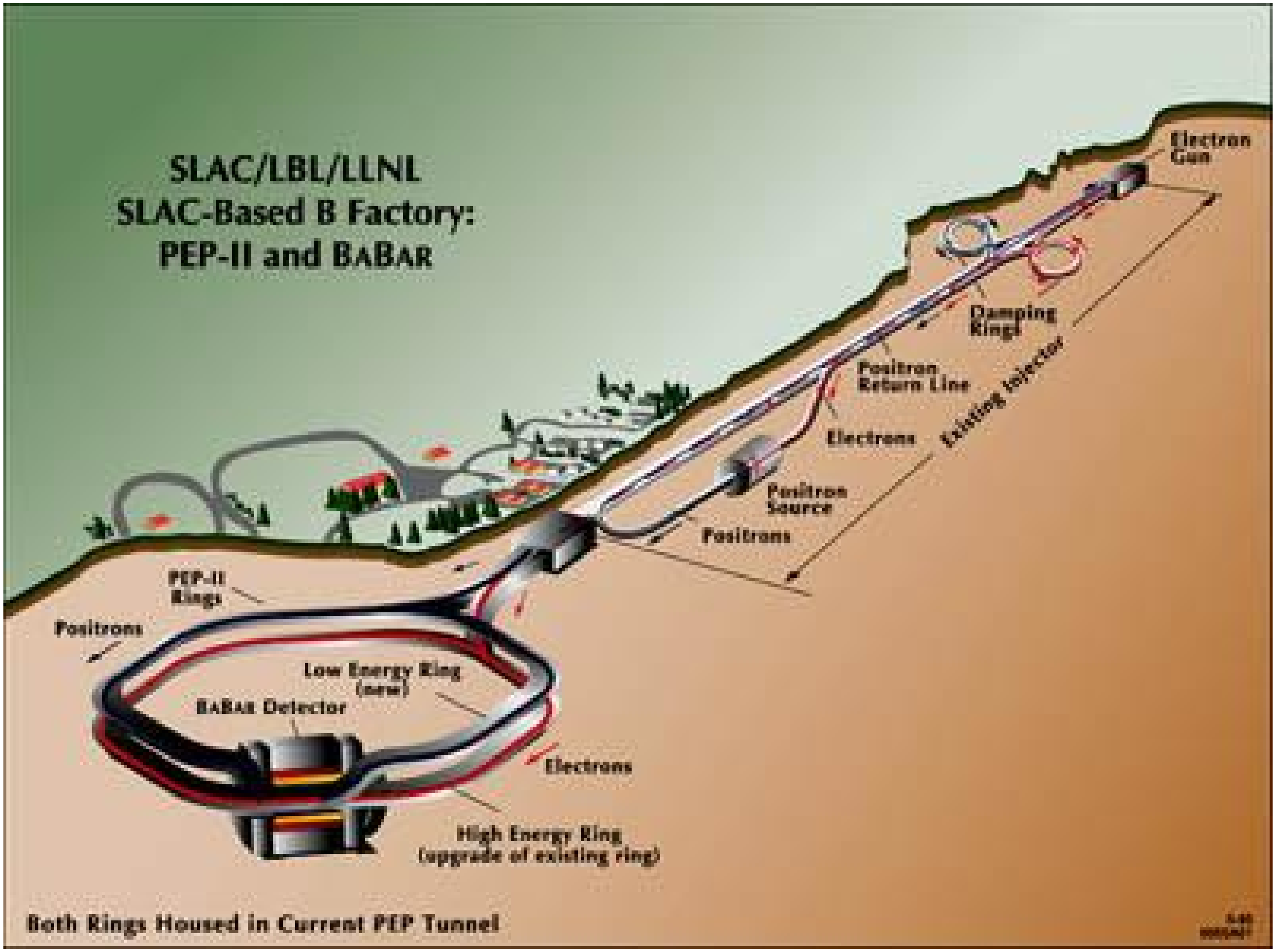
Stanford Linear Accelerator Center

- DOE laboratory managed by Stanford University
- Established 1962
- Located at Stanford University
  in Menlo Park, CA
- "Mission is to design, construct and operate state-of-the-art electron accelerators and related experimental facilities for use in high-energy physics and synchrotron radiation research."
- 3 kilometers long (e-gun to start of rings)
- 3 Nobel Prize Winners
- Home of the first U.S. Website

Stanford Linear Accelerator Center

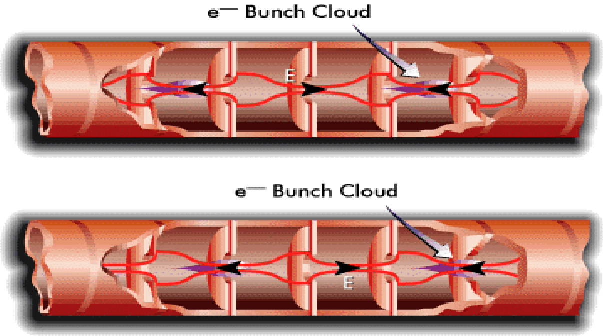
Courtesy Stanford Linear Accelerator Center

Stanford Linear Accelerator Center

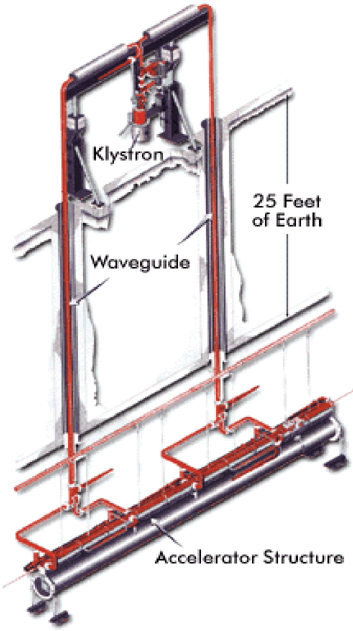


Courtesy Stanford Linear Accelerator Center

Stanford e+e- Linear Collider (SLC)



1/20,000,000,000 second later
(notice how far the bunches have moved)



Advanced Computations Department

ACD Organization/Collaborators

Advanced Computations Department

Accelerator Modeling

V. Ivanov, A. Kabel, K. Ko, Z. Li, C. Ng, L. Stingelin (PSI)

Computational Mathematics

Y. Liu, I. Malik, W. Mi, J. Scoville, K. Shah, Y. Sun (Stanford)

Computing Technologies

N. Folwell, L. Ge, A. Guetz, R. Lee, M. Wolf

Collaborators

LBNL (SCG)

E. Ng, P. Husbands, S. Li, A. Pinar

Stanford (SCCM)

G. Golub, O. Livne,

Sandia

P. Knupp, T. Tautges, L. Freitag, K. Devine

UCD (VGRG)

K. Ma, G. Schussman

LLNL (CASC)

D. Brown, K. Chand, B. Henshaw

RPI (SCOREC)

M. Shephard, Y. Luo

ACD

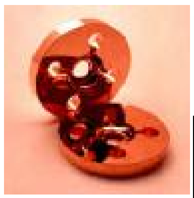
- Accurate modeling essential for modern accelerator design.
- Uncertainty in design greatly increases cost
- Accurate computer models reduce design costs and design cycle
- Need for accurate cavity design tools
- ACD develops these simulation codes
  - E&M field, resonant frequency, particle tracking simulations
  - Conformal meshes (Unstructured grid)
  - Parallel processing
- Codes: Omega3P, Tau3P, Track3P, S3P, Phi3P

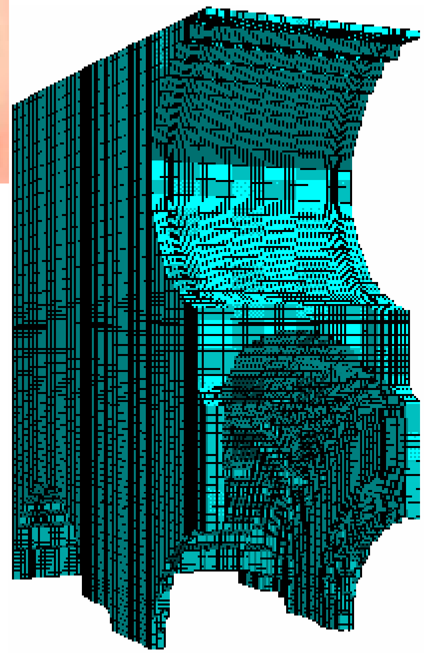
Challenges in E&M Modeling of Accelerators

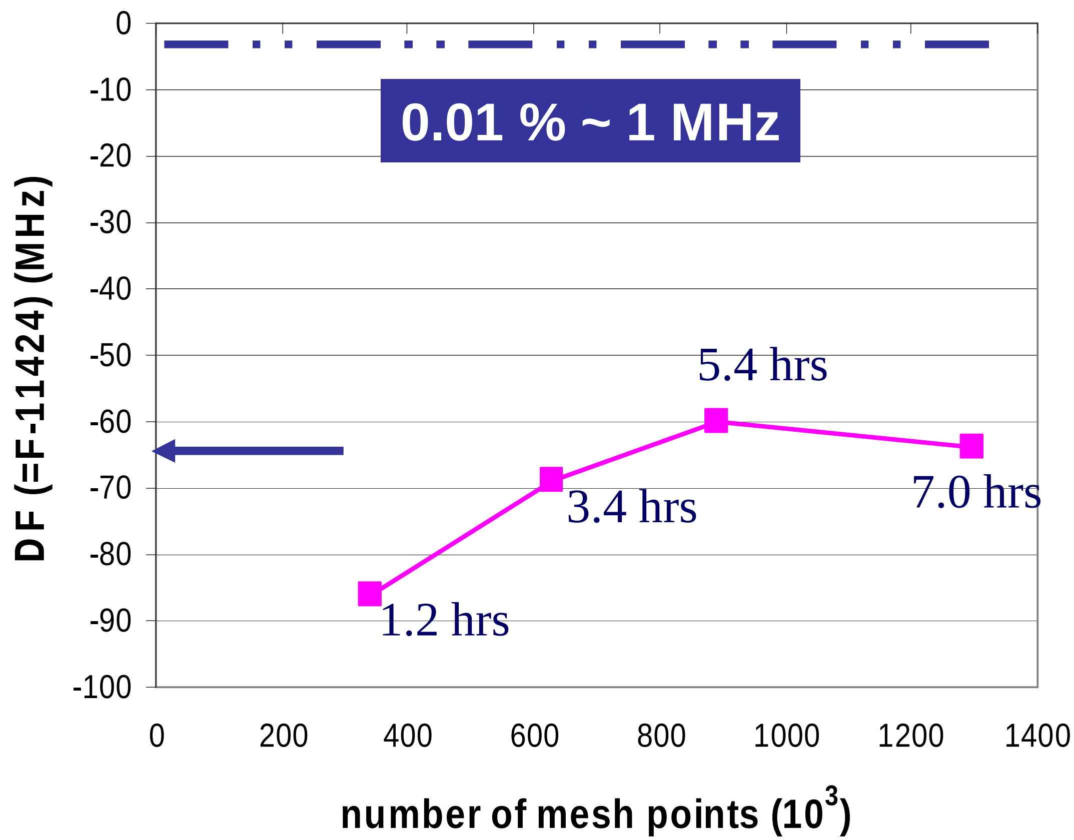
- Accurate geometry is important due to tight tolerance
  - needs unstructured grid to conform to curved surfaces
- Large, complex electromagnetic structures
  - large matrices after discretization (100's of millions of DOF), needs parallel computing for both problem storage and reduction of simulation time
- Small beam size ~ delta function excitation in time & space
  - > *Time domain* – *needs to resolve beam size leading to huge number of grid points, long run time & numerical stability issues*
  - > *Frequency domain* – *wide, dense spectrum to solve for thousands of modes to high accuracy*

Motivation for New EM Capability

Modeling RDDS Cell with standard accelerator code MAFIA using Structured Grid on Desktops demonstrates the need for MORE ACCURATE EM codes



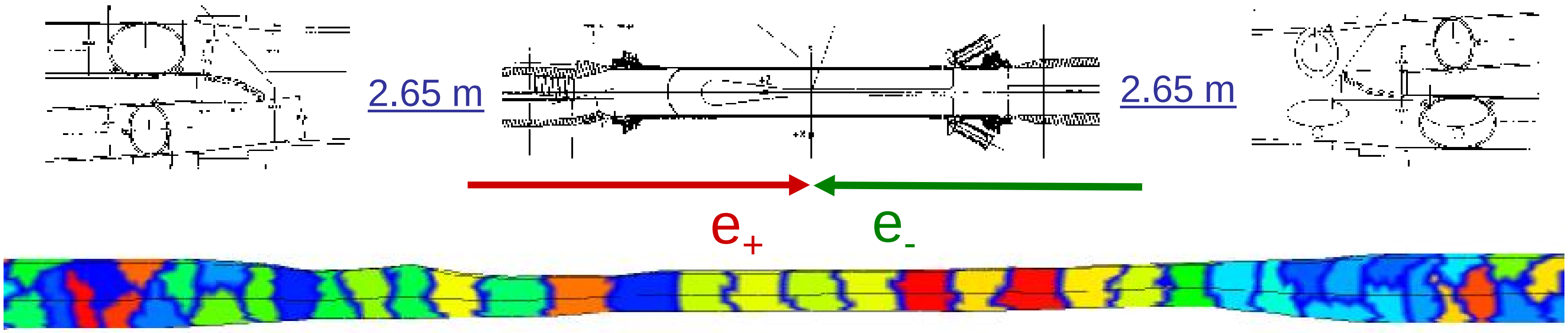


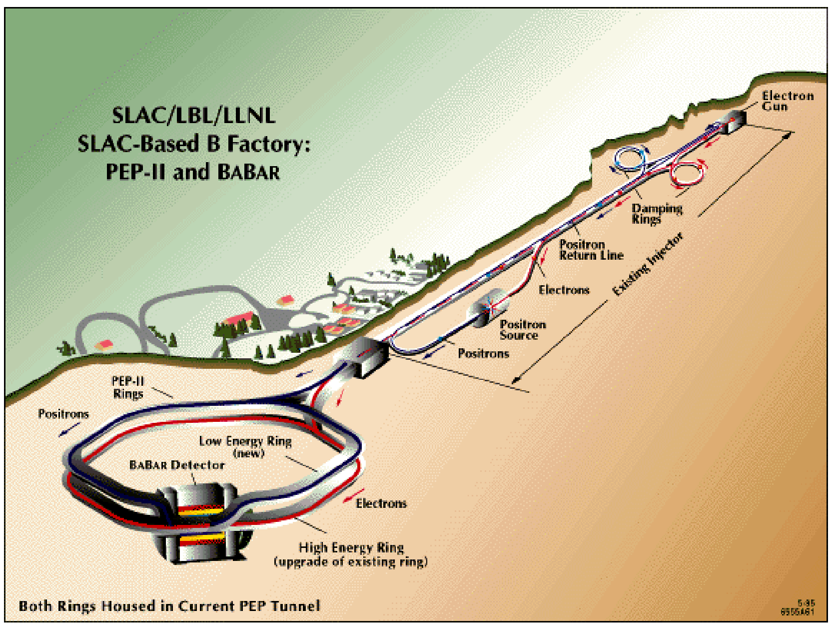


Typical ACD problems

Heating in PEP-II Interaction Region

Left crotch *Center beam pipe* *Right crotch*

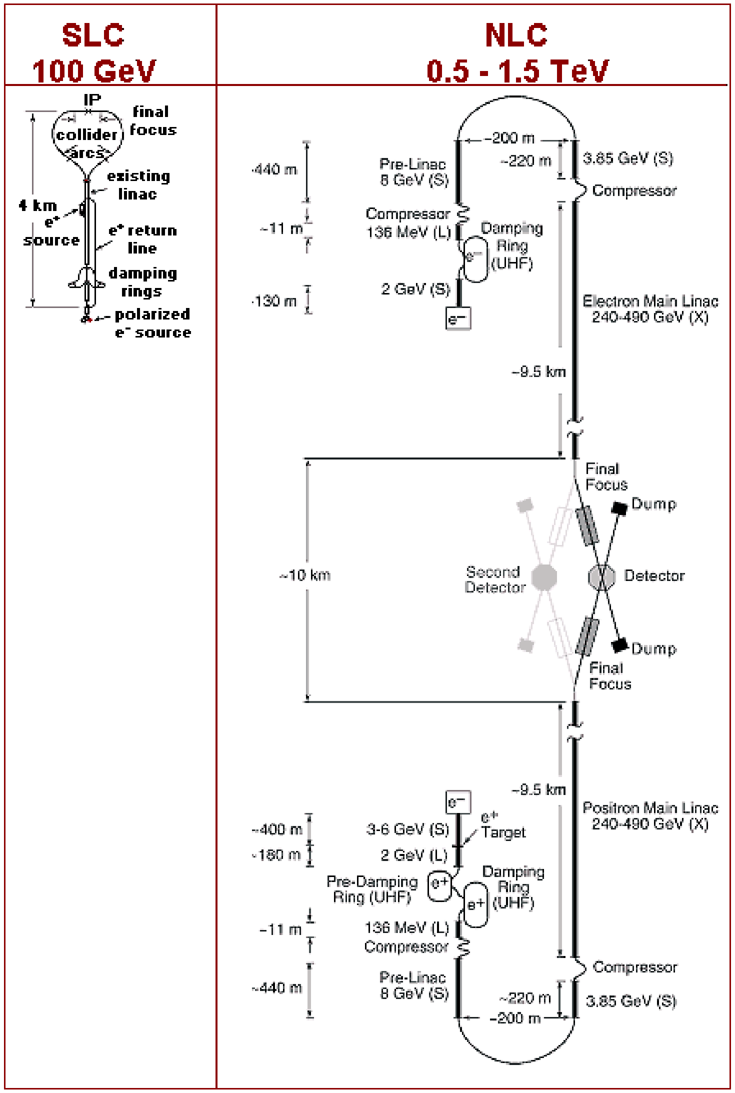




FULL-SCALE OMEGA3P MODEL FROM CROTCH TO CROTCH

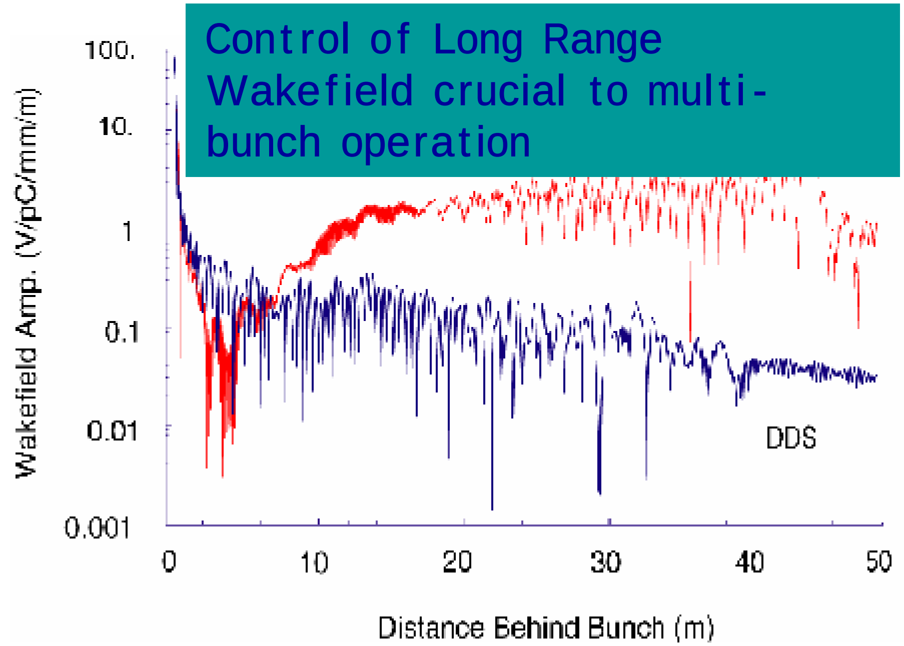
Beam heating in the beamline complex near the IR limited the PEP-II from operating at high currents. Omega3P analysis helped in redesigning the IR for the upgrade.

Next Linear Collider (NLC)



	SLC	NLC
Center of Mass	100 GeV	500 GeV
Bunches per pulse	1	95
Operating Frequency	(S) 2.856 GHz	(X) 11.424 GHz
Number of Cavities	80,000	2 million
Post-Tuning	yes	No

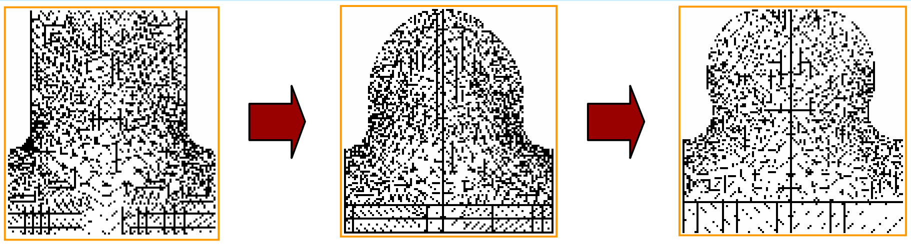
Control of Long Range Wakefield crucial to multi-bunch operation



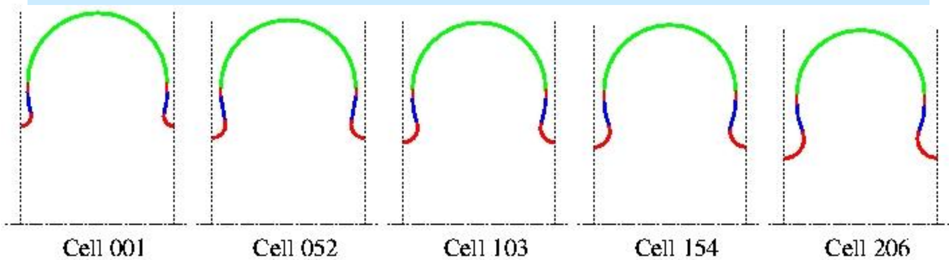
The NLC Accelerating Structure

206-Cell Round Damped Detuned Structure RDDS

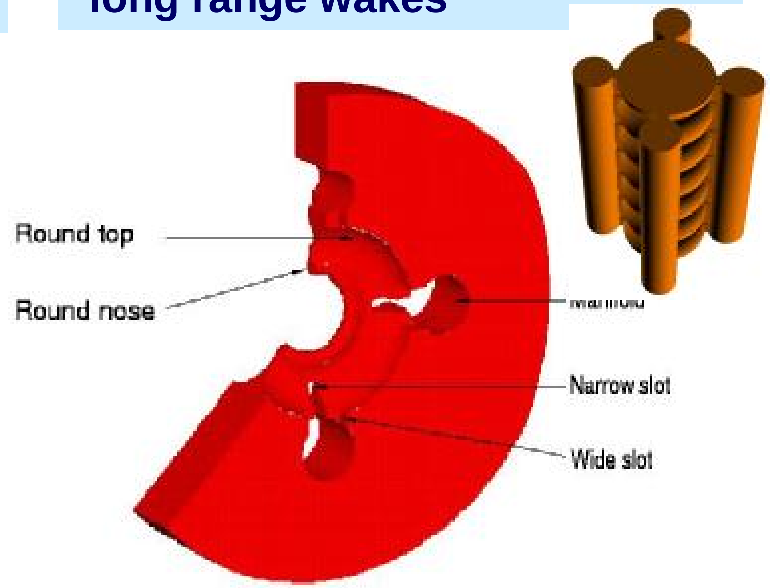
Cell optimized to increase shunt impedance (~14%) & minimize surface gradients



Cell to cell variation of order microns to suppress short range wakes by detuning



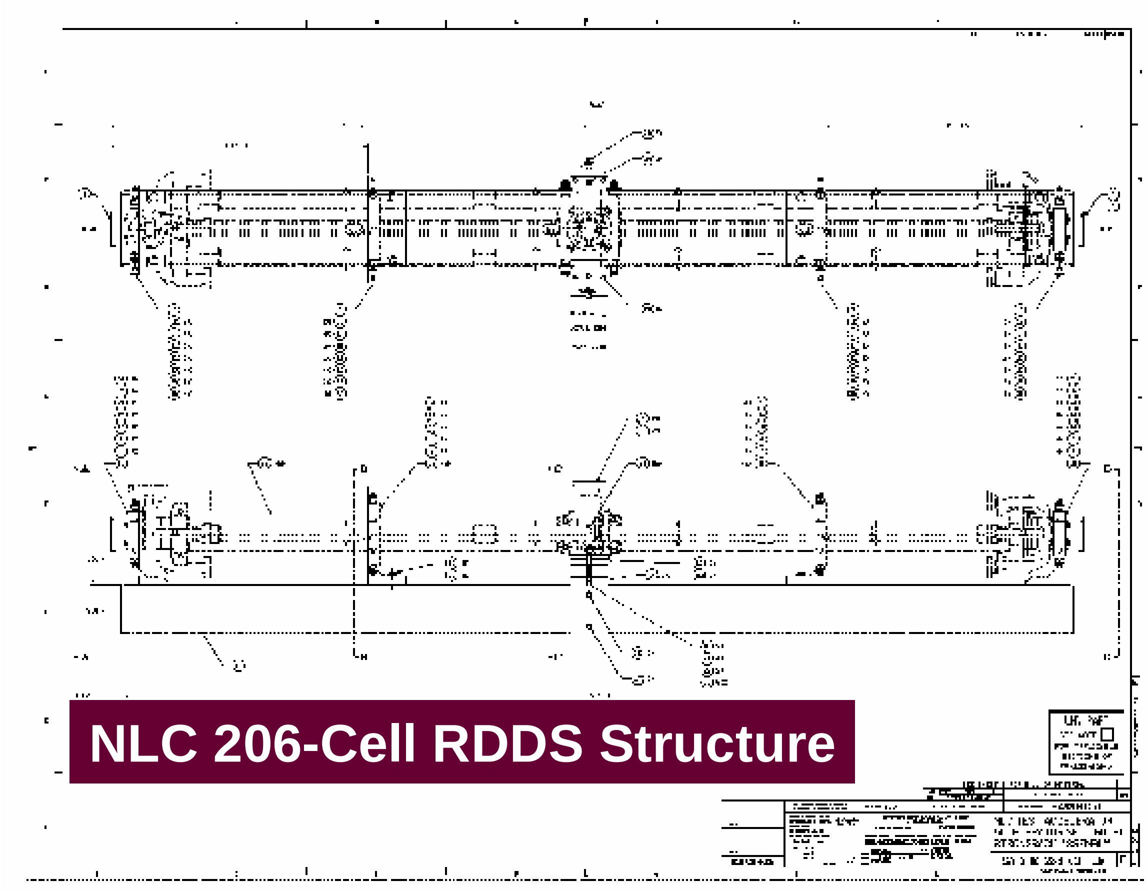
Manifold damping to suppress long range wakes



11 cavity dimensions

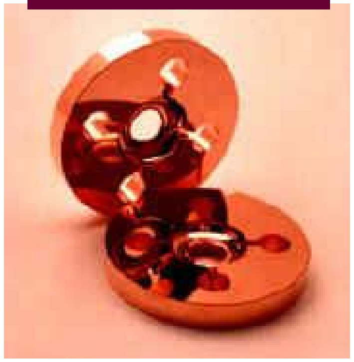
- Ø Needs accelerating frequency calculated to 0.01% accuracy to maintain structure efficiency
- Ø Optimized design could save $100 million in machine cost

NLC Structure Design Requirements

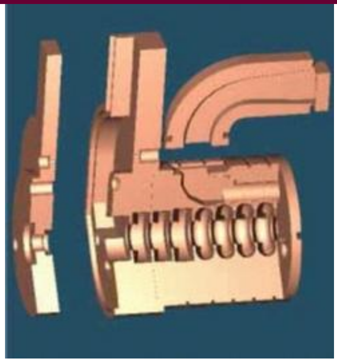


NLC 206-Cell RDDS Structure

RDDS Cell

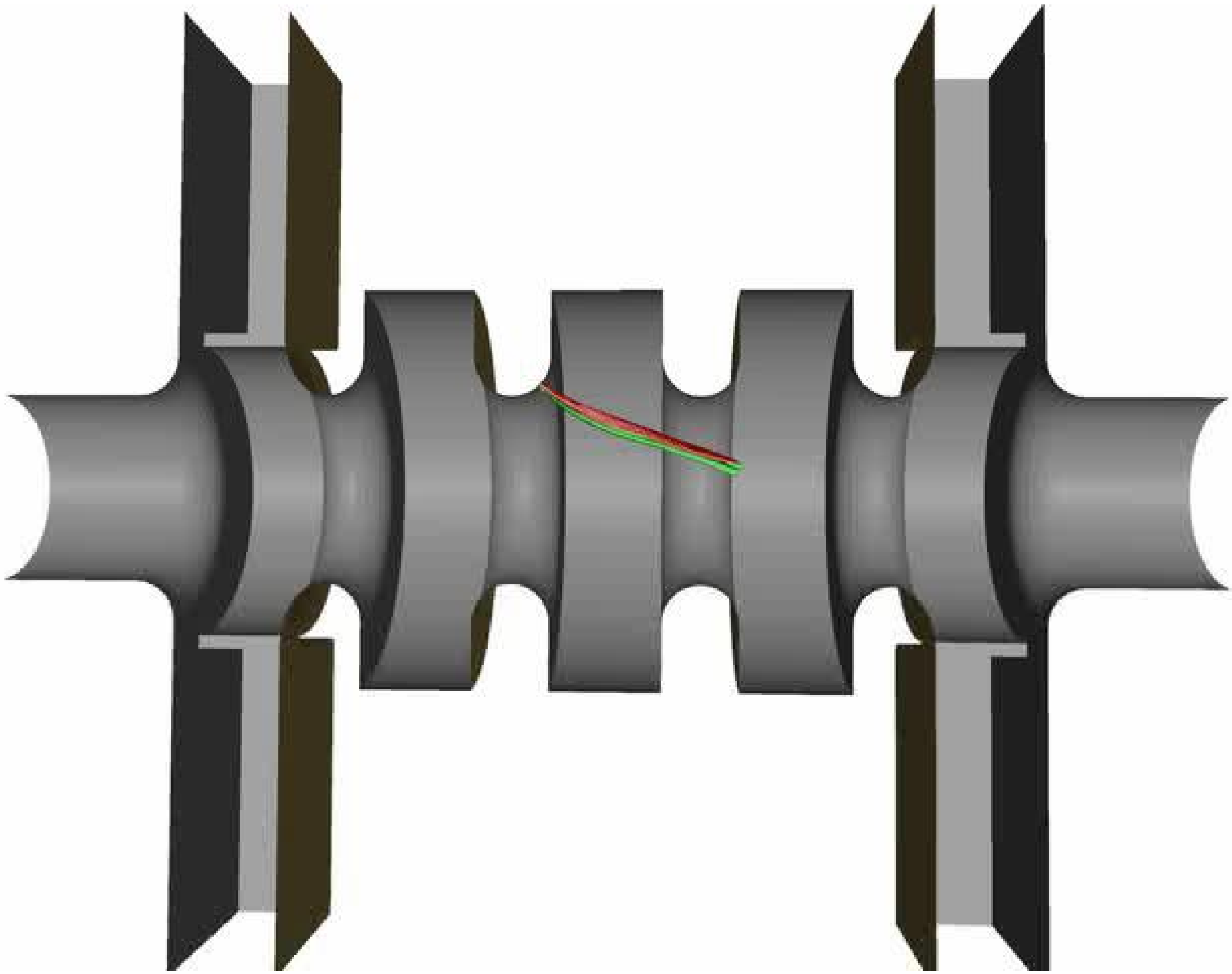


RDDS Section



- RDDS Cell : Design to 0.01% accuracy in accelerating frequency,
- RDDS Section : Model damping/detuning of dipole wakefields.

Particle Tracking in 5 Cell RDDS (Tau3P/Track3P)



Cyclotron COMET (Omega3P)

First ever detailed analysis of an entire cyclotron structure
- L. Stingelin, PSI

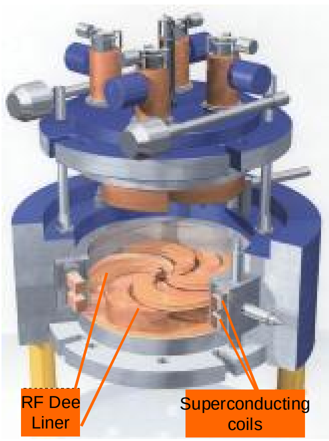
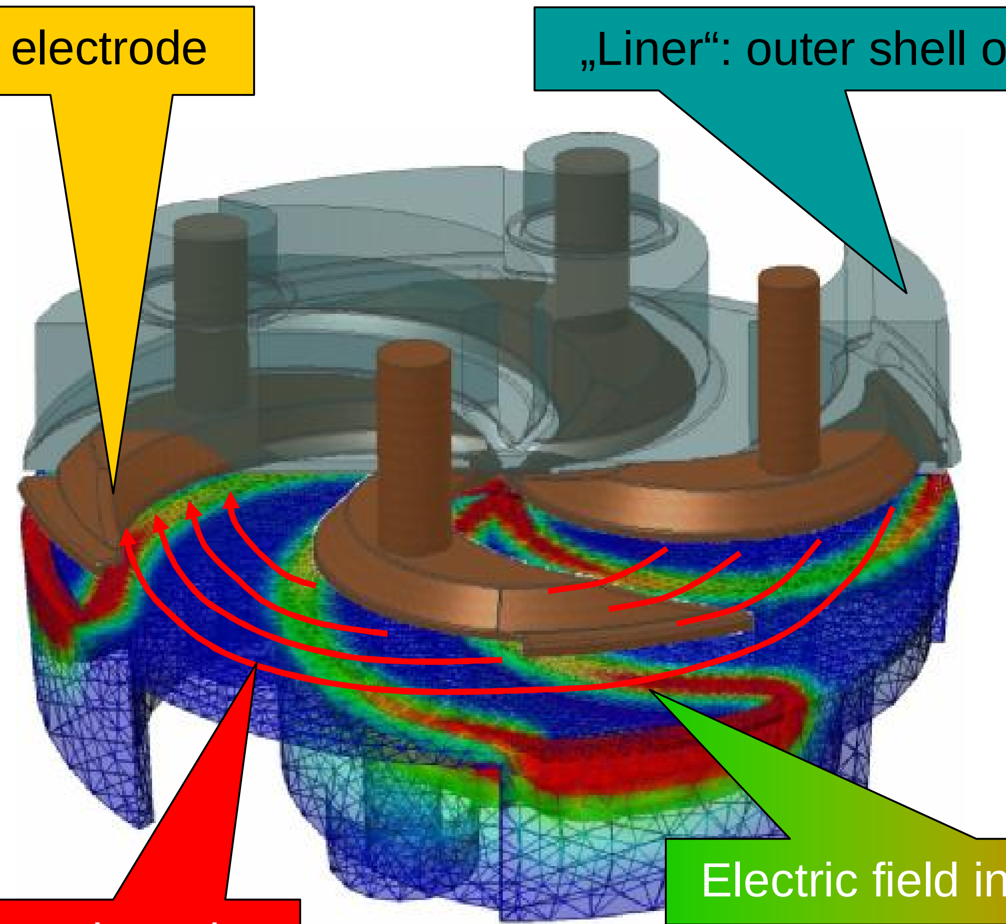
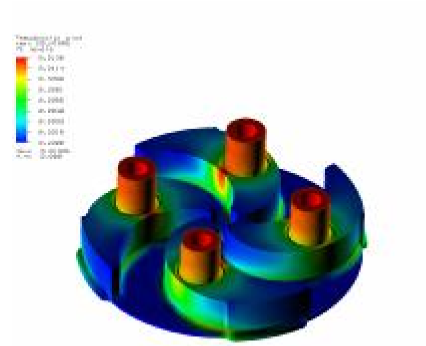


Fig 1. The cyclotron COMET





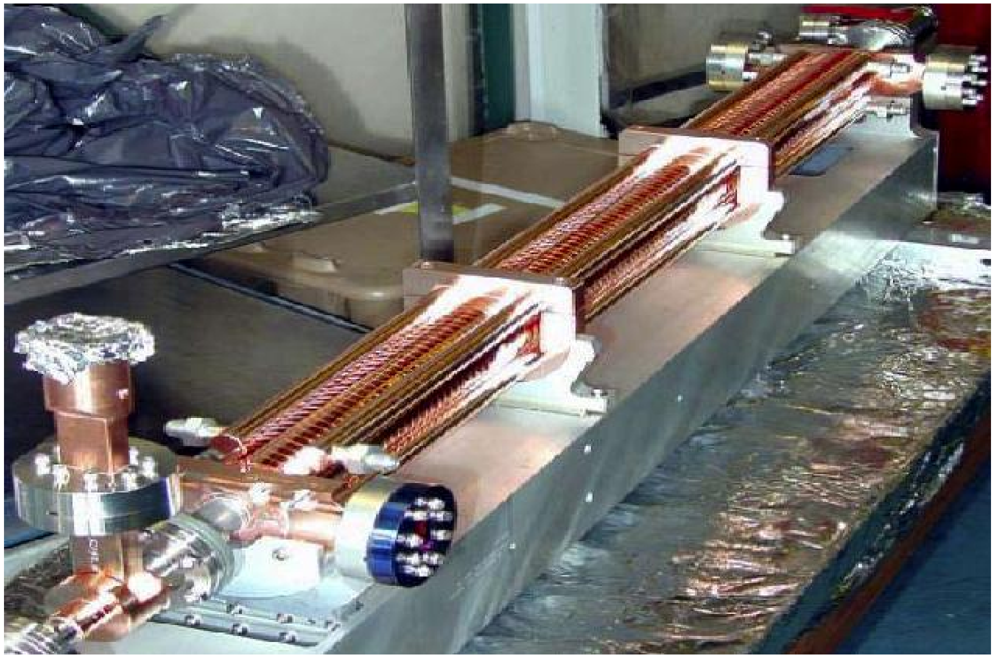
Magnetic Field

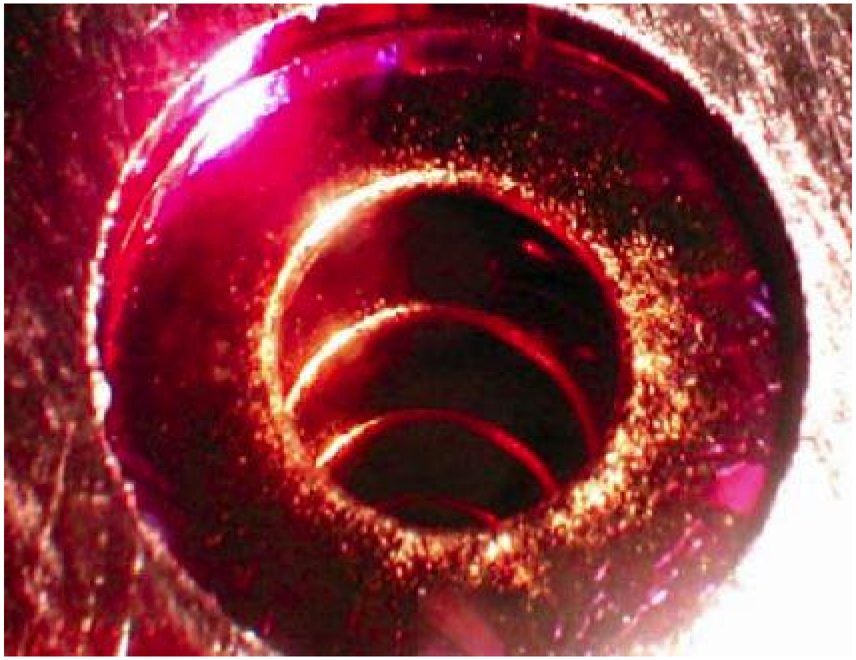
Where we want to go

End-to-end NLC Structure Simulation

(J. Wang, C. Adolphsen – SLAC)

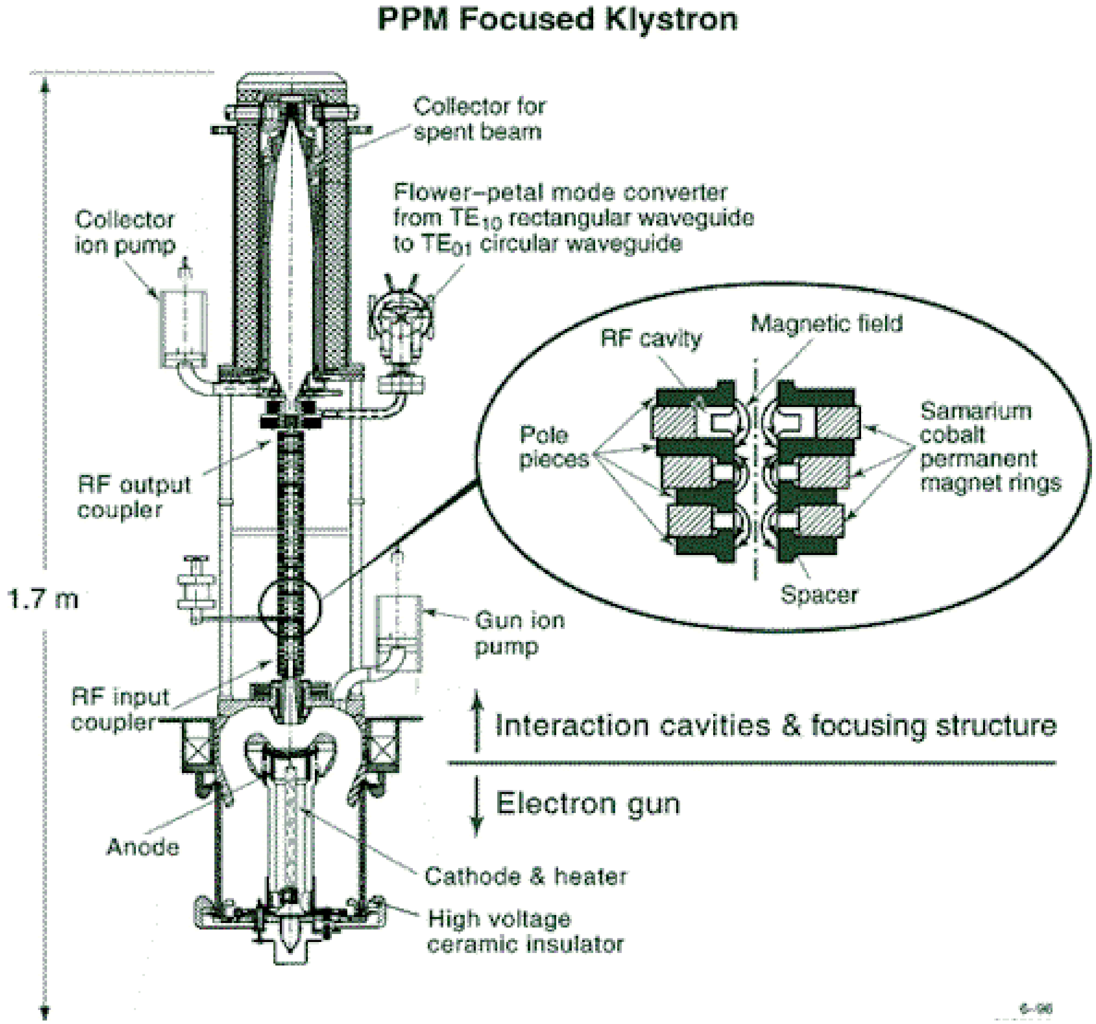
- NLC X-band structure showing damage in the structure cells after high power test
- Theoretical understanding of underlying processes lacking so realistic simulation is needed





End-to-end NLC Klystron Simulation

Field and particle data estimated to be TB size



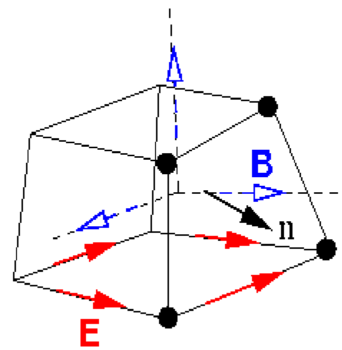
Tau3P
Parallel Time-Domain 3D Field Solver

Parallel Time-Domain Field Solver – Tau3P

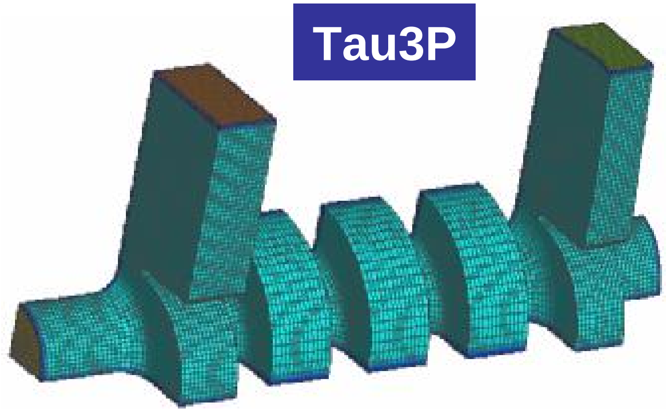
Follows evolution of E and H fields on primary/dual meshes (hexahedral, tetrahedral and pentahedral elements) using leap-frog scheme in time (DSI scheme)

$$\oint E \bullet ds = -\iint \frac{\partial B}{\partial t} \bullet dA$$

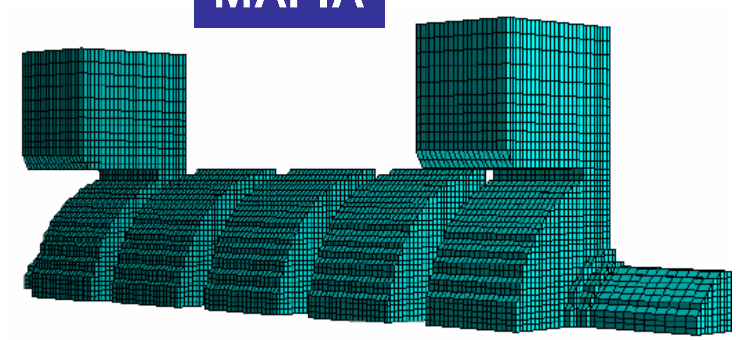
$$\oint H \bullet ds^* = \iint \frac{\partial D}{\partial t} \bullet dA^* + \iint j \bullet dA^*$$



Tau3P



MAFIA



Discrete Surface Integral Method

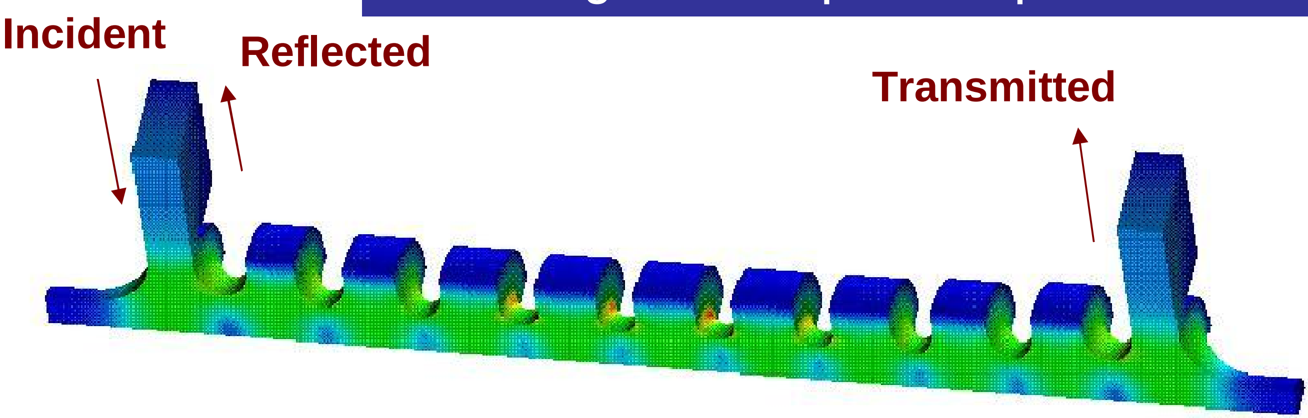
- Electric fields on primary grid
- Magnetic fields on dual grid
- Primary and dual grids non-orthogonal
- Dual grid constructed by joining centers of primary cells
- Electric and magnetic fields advanced in time using the leapfrog algorithm
- Reduces to conventional finite difference time domain method (FDTD) for non-orthogonal grids
- Conforms to complicated geometry by appropriate choice of element types

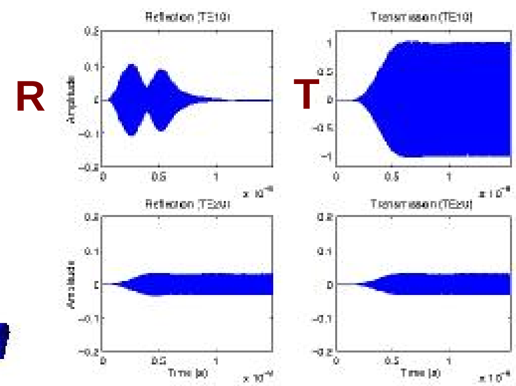
Ref: N. K. Madsen, J. Comp. Phys., **119**, 34 (1995)

Tau3P Applications

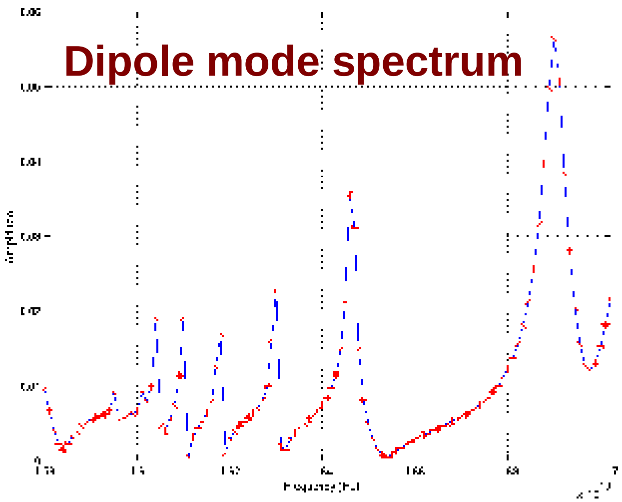
Time Domain Design & Analysis (Tau3P)

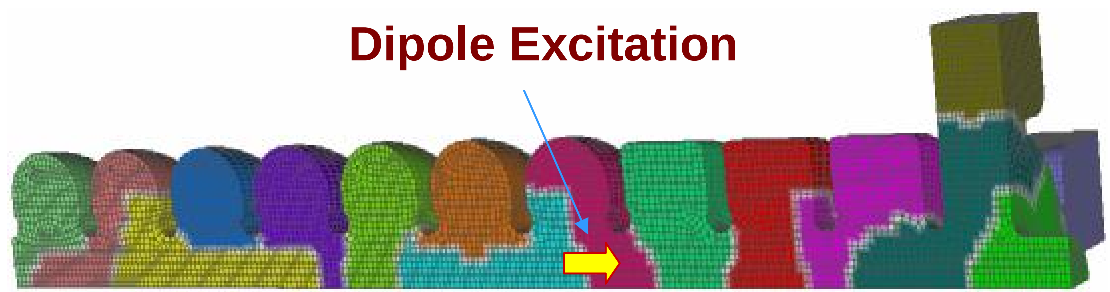
Matching NLC Input Coupler w/ Inline Taper





Output Coupler loading on HOM modes at the RDDS output end

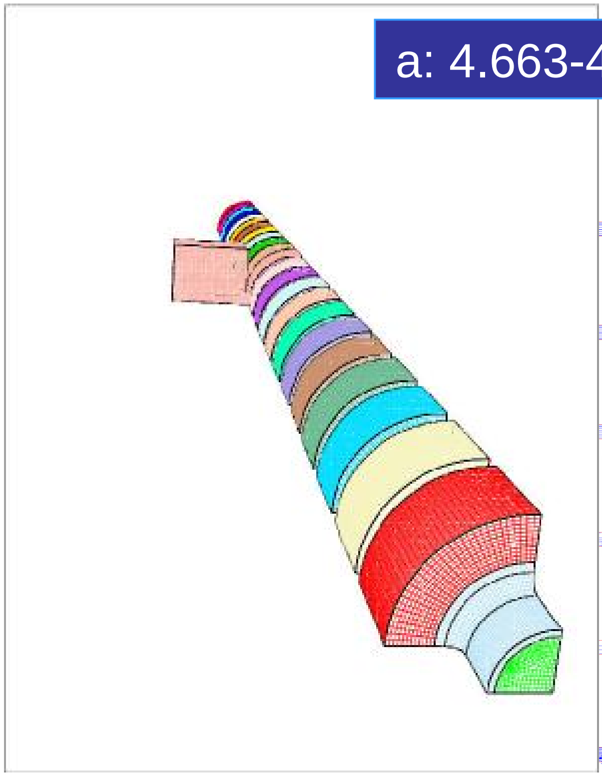


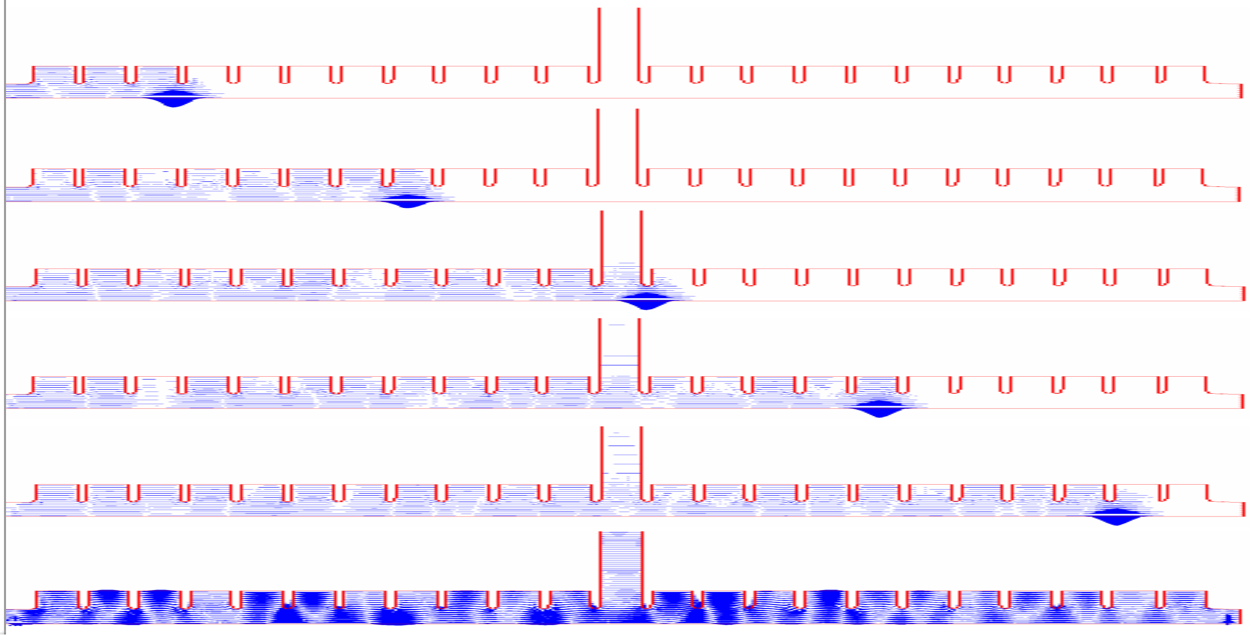


Wakefield **Calculation (Tau3P)**

- Response of a 23-cell X-Band Standing Wave Structure w/ Input Coupler & Tapered Cells to a transit beam in Tau3P.
- Direct wakefield simulation of exact structure to verify approximate results from circuit model.

a: 4.663-4.875 mm, b:10.796-10.879 mm, t: 2.541-2.684 mm



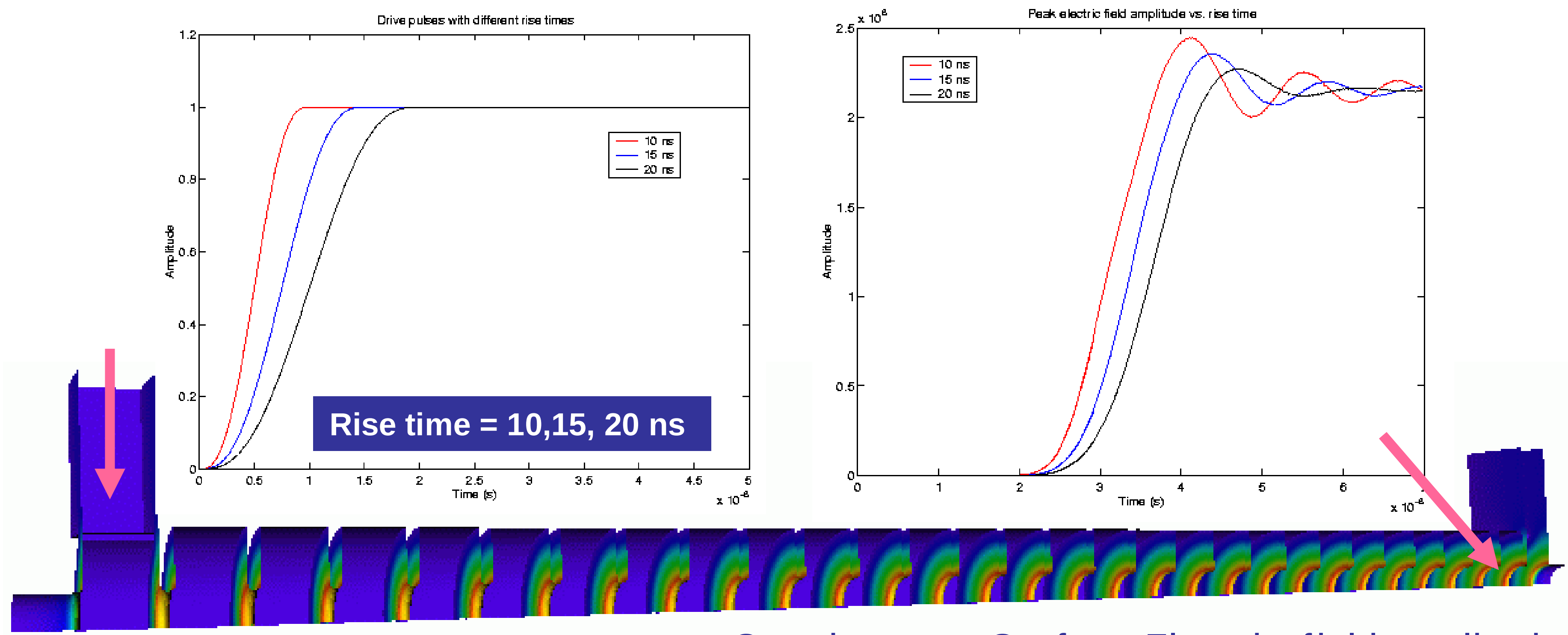


Determining Peak Fields (Tau3P)

- When and where Peak Fields occur during the pulse?
- Transient fields 20% higher than steady-state value due to dispersive effects

Drive pulse

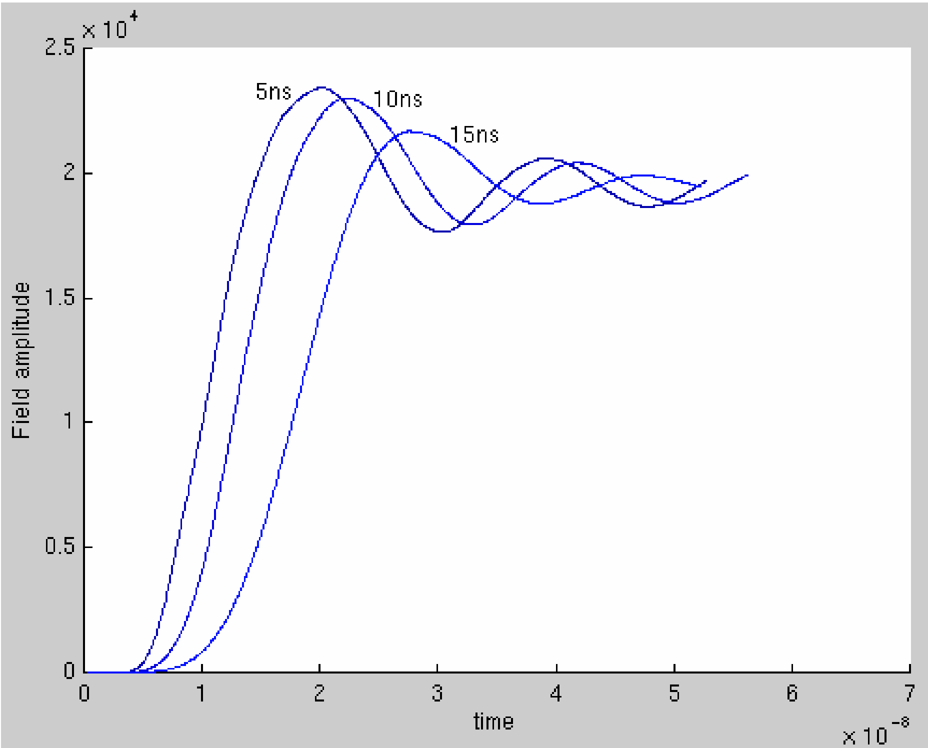
Electric field vs time



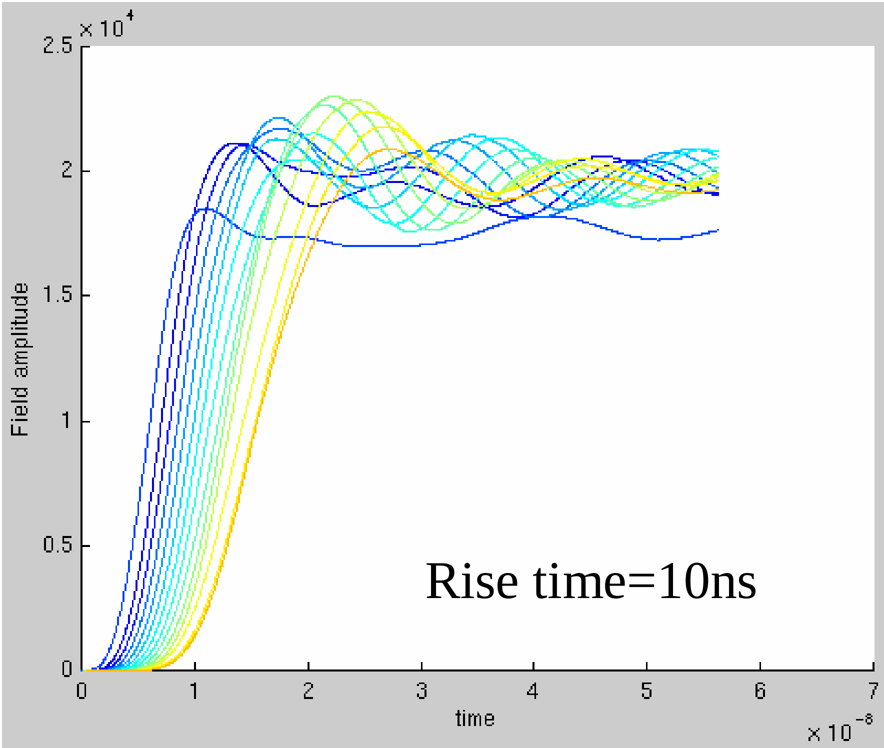
Steady-state Surface Electric field amplitude

15-Cell H90VG5 Model

- Peak field appears near the middle of the structure
- About 25% overshoot in peak field due to the narrower bandwidth

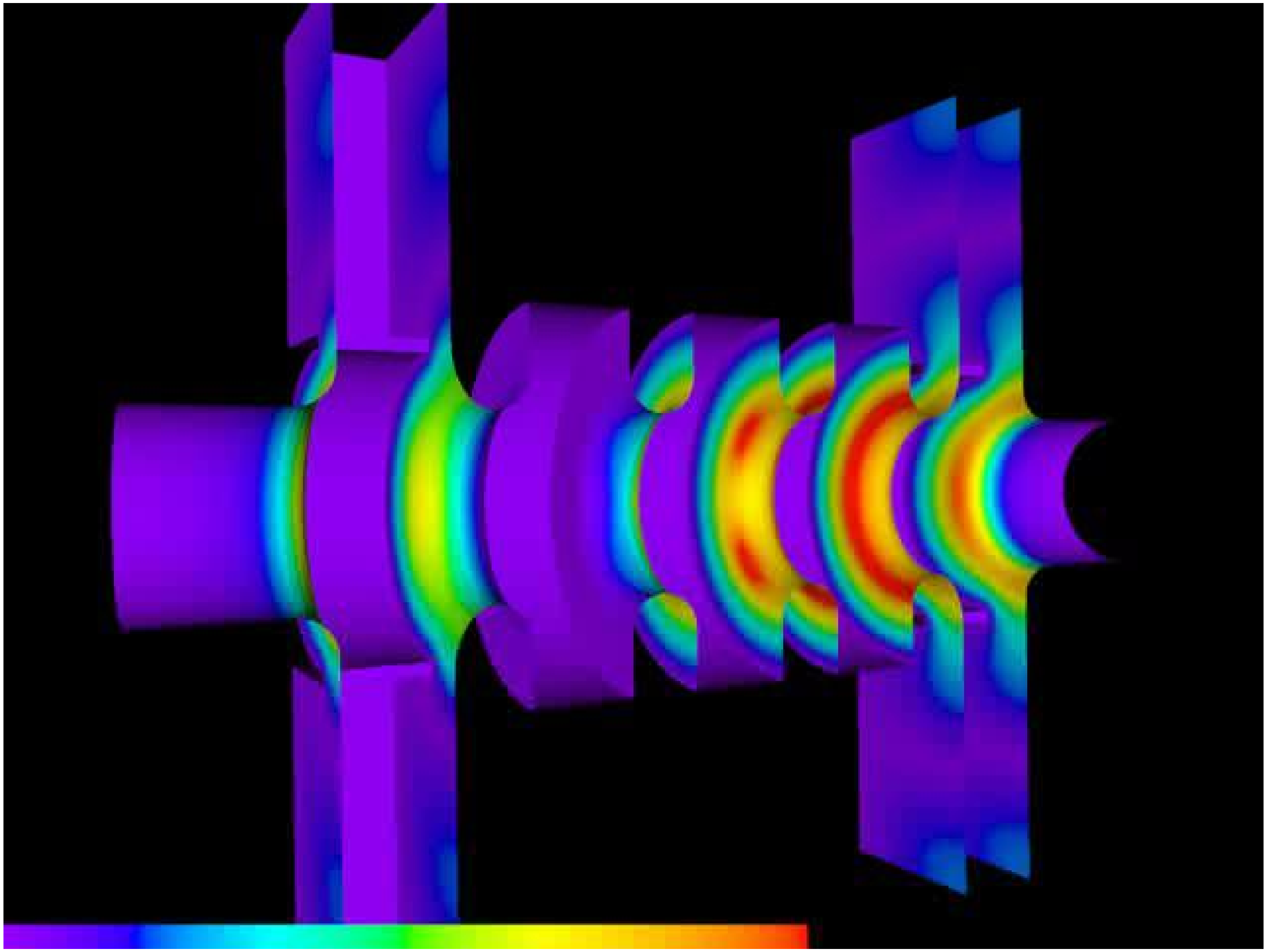


Overshoots for different rise times



Fields at different cell disks as a function of time

Electric Fields (Tau3P)

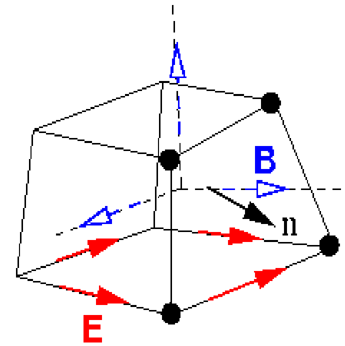


Tau3P Matrices

Discrete Surface Intergral Formulation

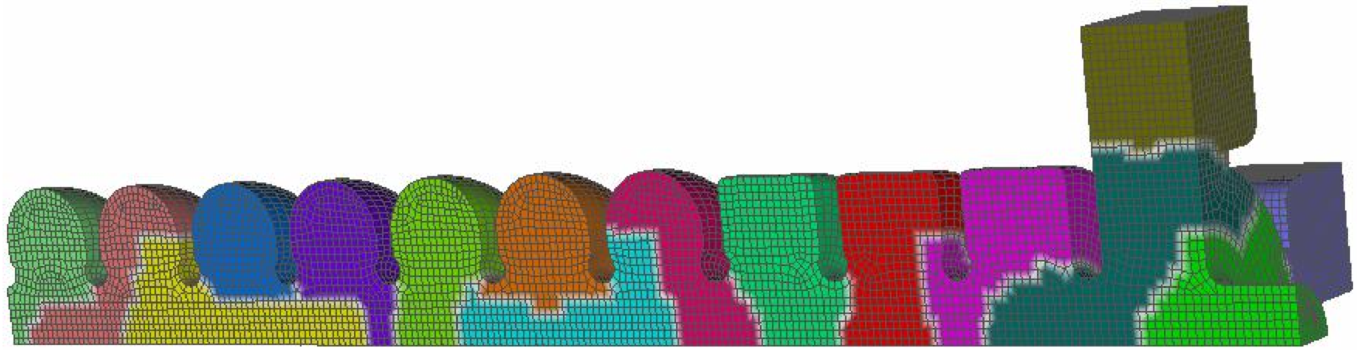
$$\oint E \bullet ds = -\iint \frac{\partial B}{\partial t} \bullet dA$$

$$\oint H \bullet ds^* = \iint \frac{\partial D}{\partial t} \bullet dA^* + \iint j \bullet dA^*$$

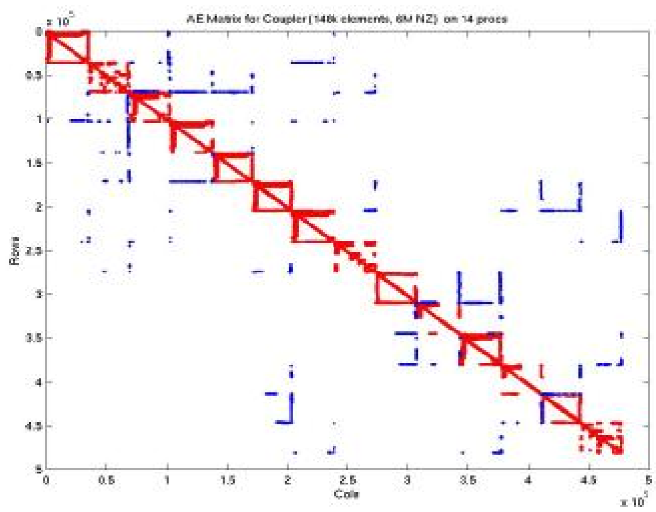


- The DSI formulation yields:
  - efield += α*AH*hfield
  - hfield += β*AE*efield
    - efield, hfield are vectors of field projections along edges/dual-edges
    - AH,AE are matrices
    - α, β are constants proportional to dt

Tau3P Implementation

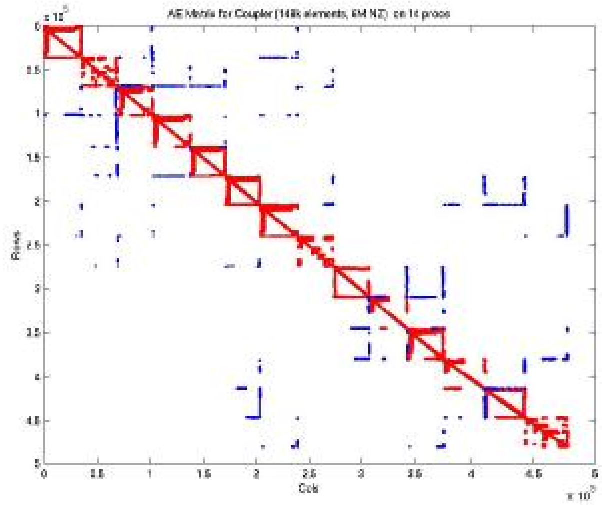


Example of Distributed Mesh



Typical AE Distributed Matrix

Tau3P Matrix Properties



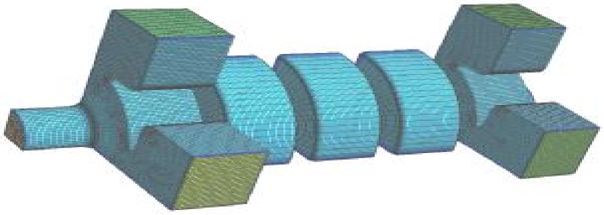
Typical Distributed Matrix

- Very Sparse Matrices
  - 4-20 nonzeros per row
- 2 Coupled Matrices (AH,AE)
- Nonsymmetric (Rectangular)

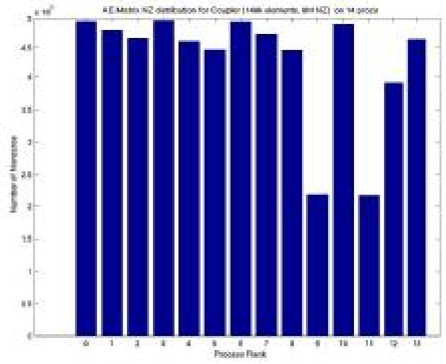
Tau3P Performance Problems

Load Balancing Issues in Tau3P

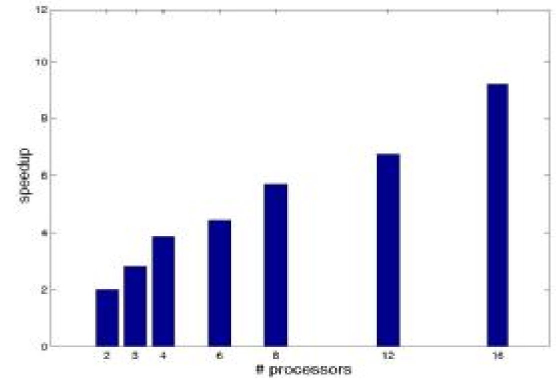
- Load balancing problem inTau3P Modeling of NLC Input Coupler.
  - Unstructured meshes lead to matrices for which nonzero entries are not evenly distributed.
  - Complicates work assignment and load balancing in a parallel setting.
  - Tau3P originally used ParMETIS to partition the domain to minimize communication.



NZ Distribution over 14 cpu's



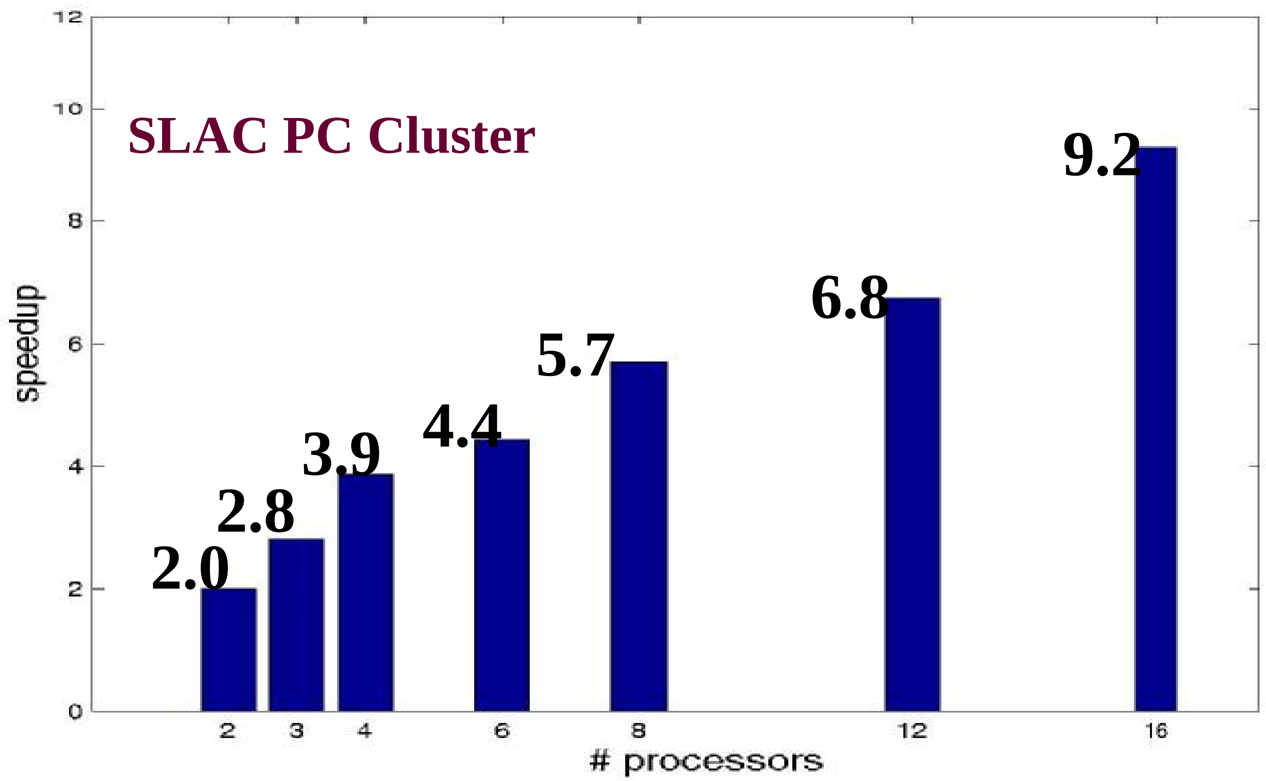
Parallel Speedup



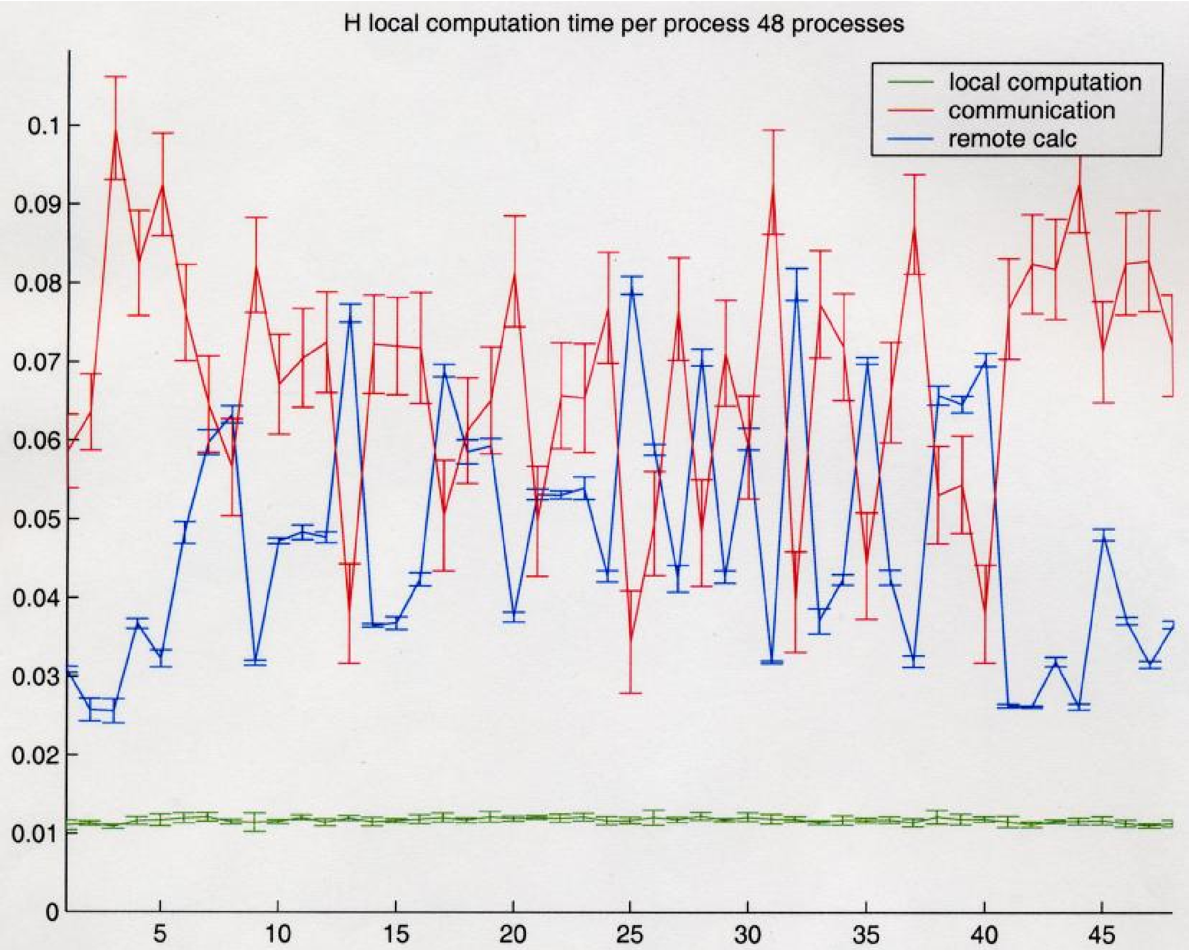
Parallel Performance of Tau3P

- 257K hexahedrons
- 11.4 million non-zeroes

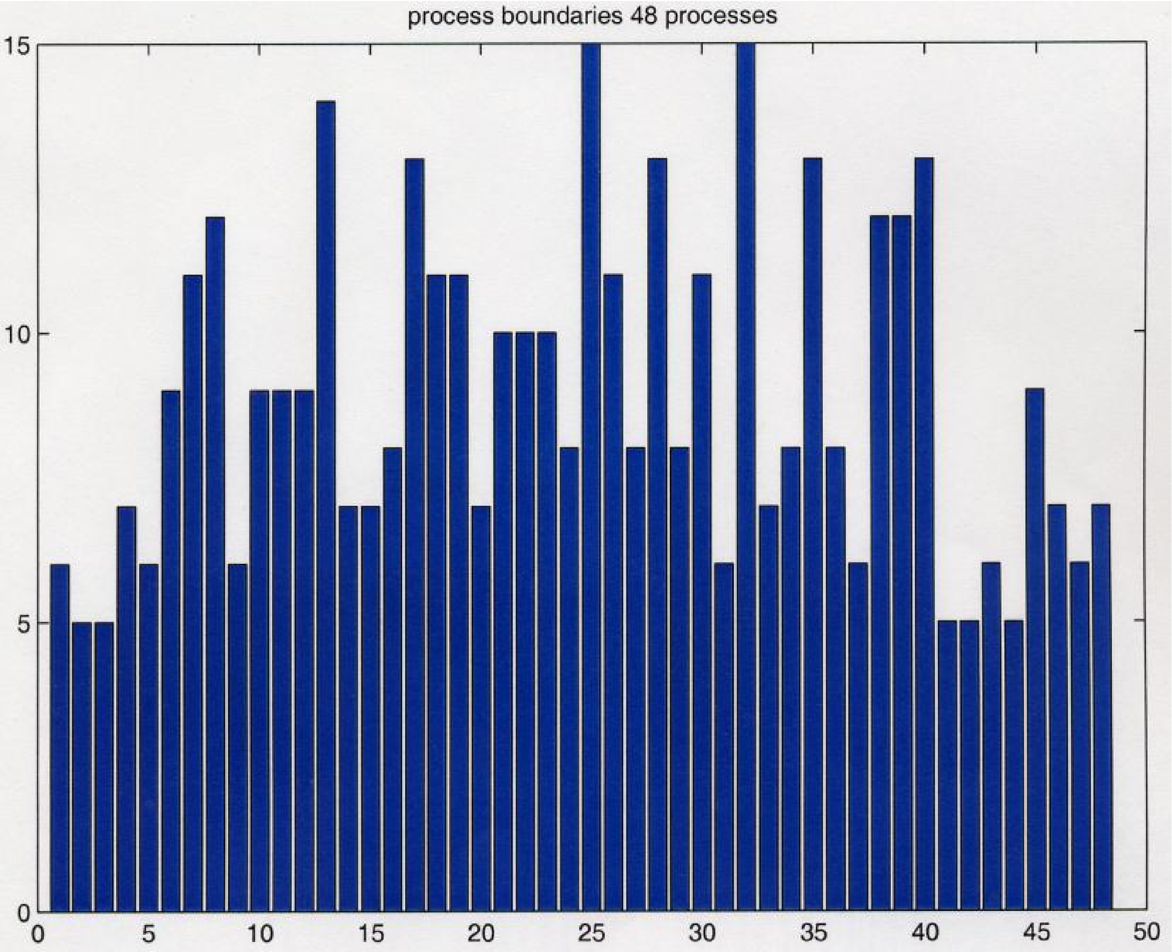
Parallel Speedup



Communication in Tau3P (ParMetis Partitioning)



Communication in Tau3P (ParMetis Partitioning)



Improved Mesh Partitioning Schemes

Luxury in Tau3P Mesh Partitioning

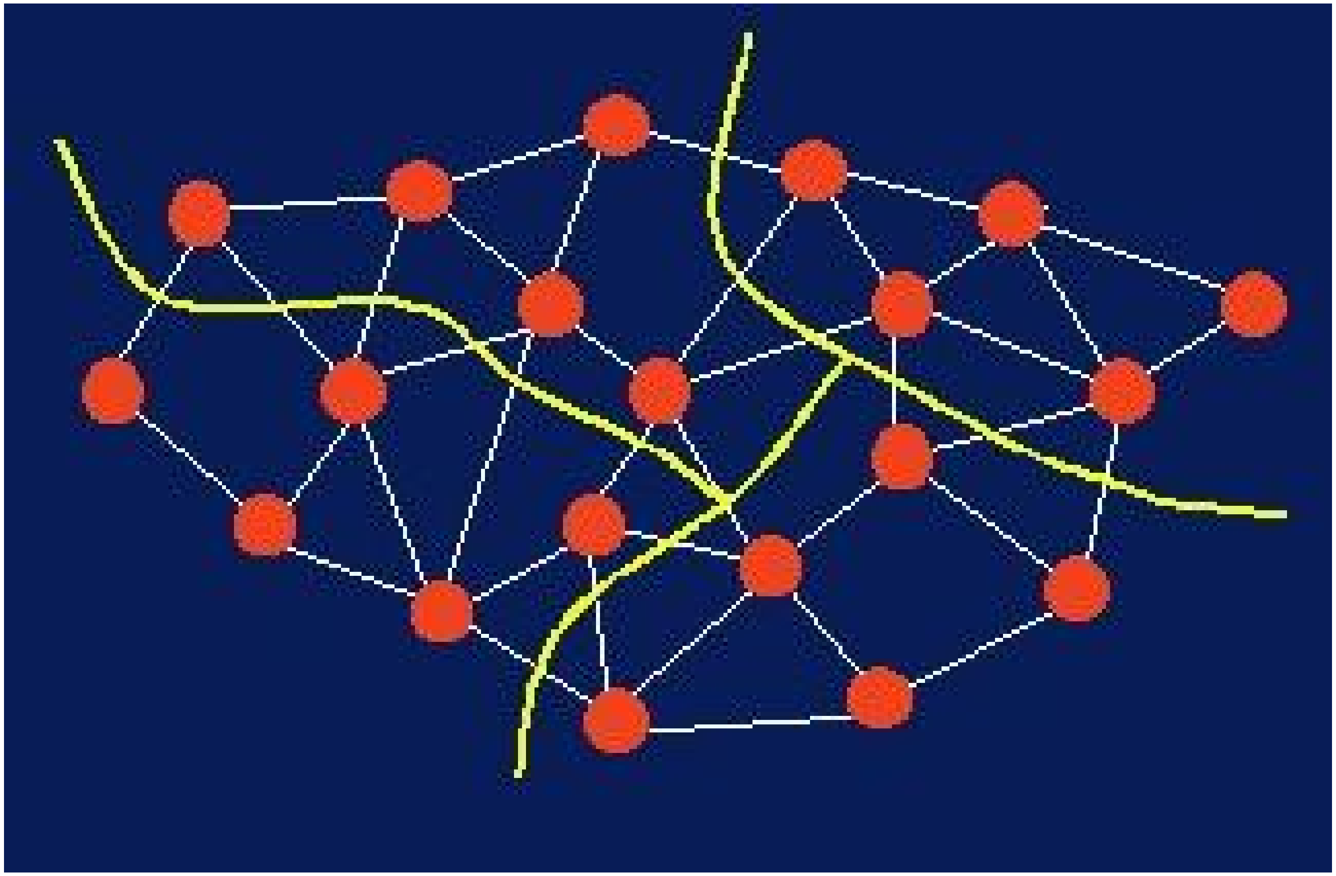
- Long simulation times
  - Tens of thousands of CPU hours
- Long time spent in time stepping
  - Millions of time steps
- Problem initialization short
- Static (not dynamic) mesh partitioning
- Willing to pay HIGH price upfront for increased performance of solver

Zoltan Overview

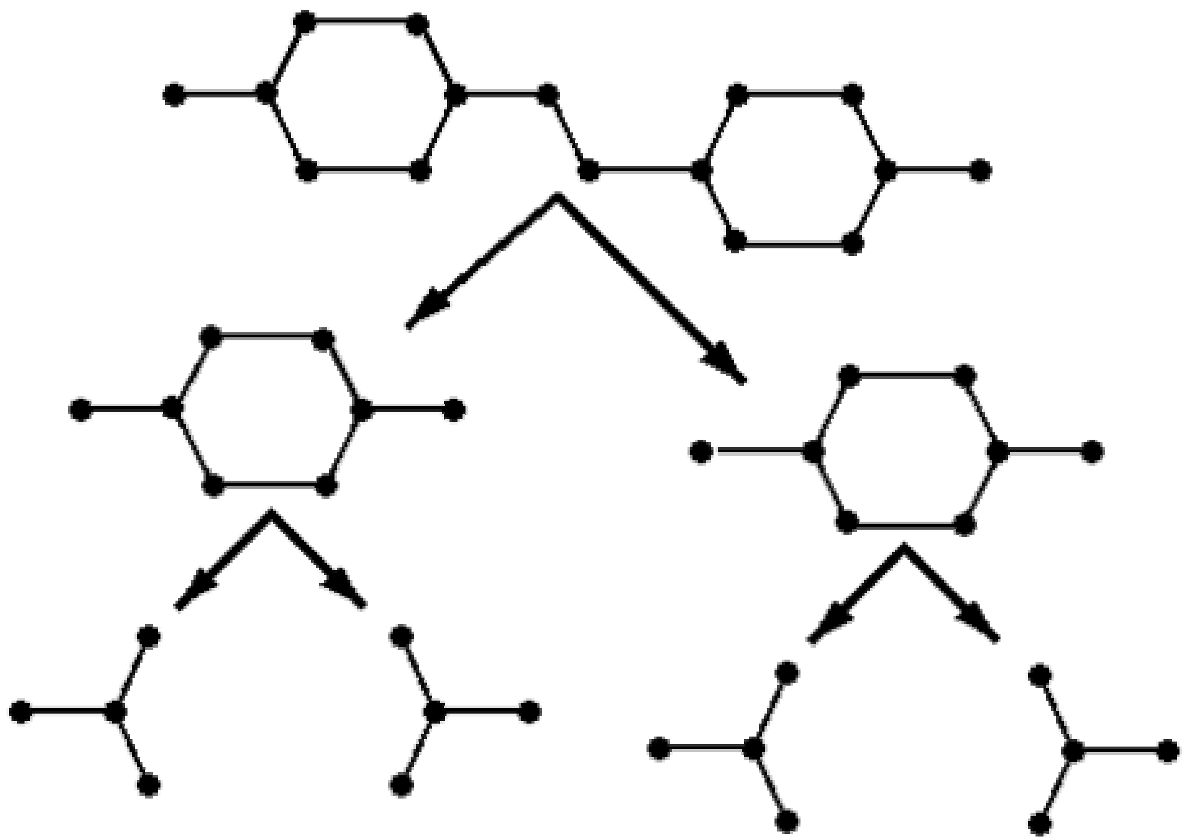
- Developed at Sandia National Laboratory (NM)
- Collection of Data Management Services for Parallel, Unstructured, Adaptive, and Dynamic Applications
- Supports Several Load Balancing Methods:
  - Graph Partitioning Algorithms
    - ParMETIS
    - Jostle
  - Geometric Partitioning Algorithms (1D/2D/3D)
    - Recursive Coordinate Bisection (RCB)
    - Recursive Inertial Bisection (RIB)
    - Hilbert Space-Filing Curve (HSFC)
    - Octree Partitioning (various traversal schemes including HSFC)
    - Refinement Tree Based Partitioning (mesh refinement)
- Supports Dynamic Load Balancing/Data Migration

Zoltan Partitioning Methods

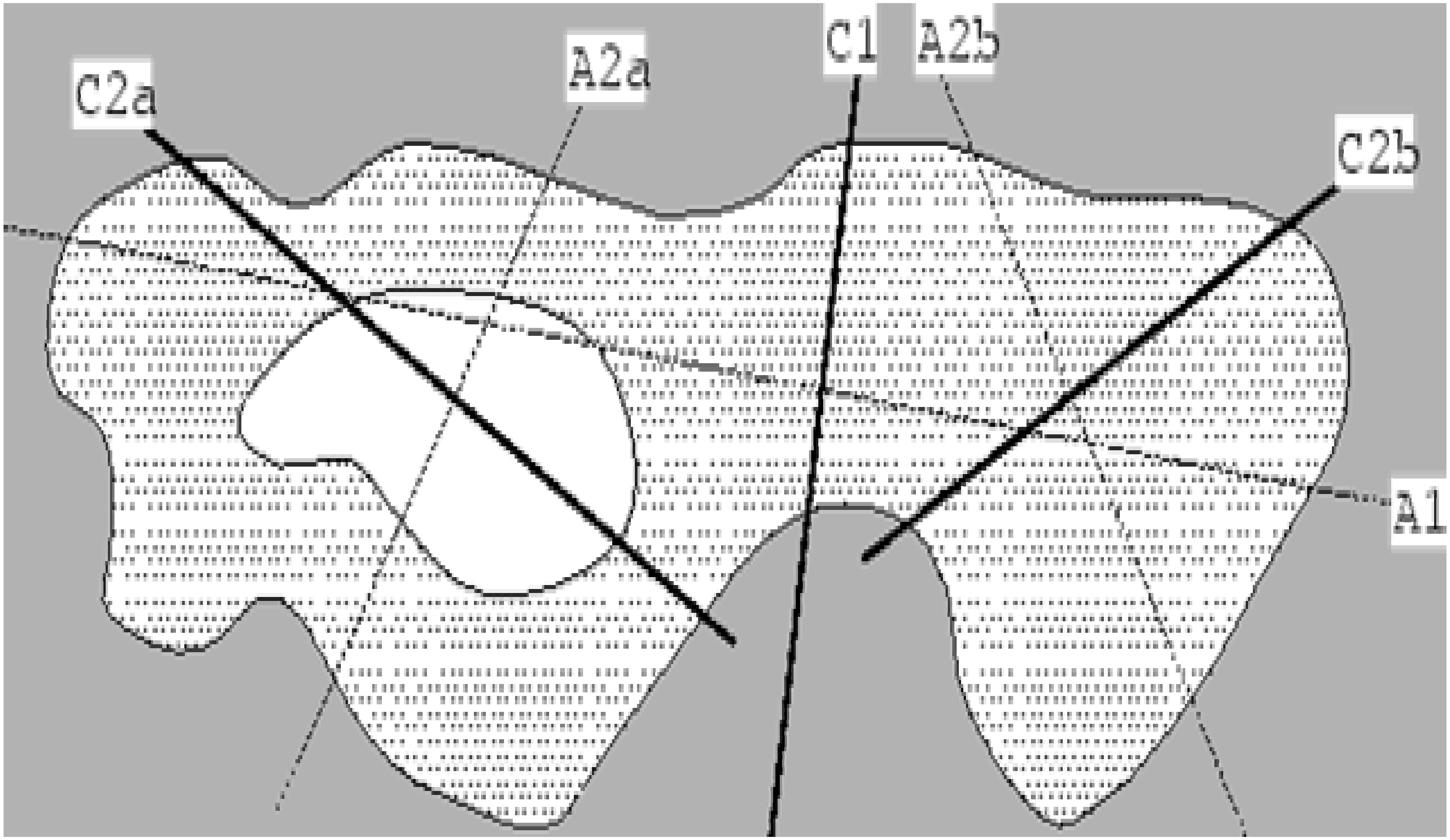
ParMETIS (Graph)



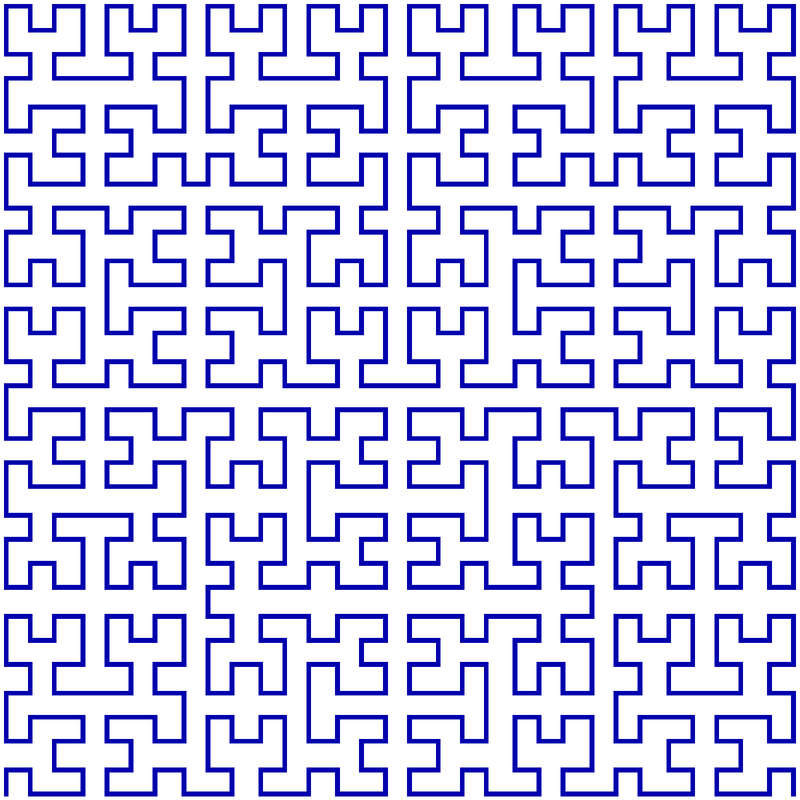
Recursive Coordinate Bisection (Geometric)



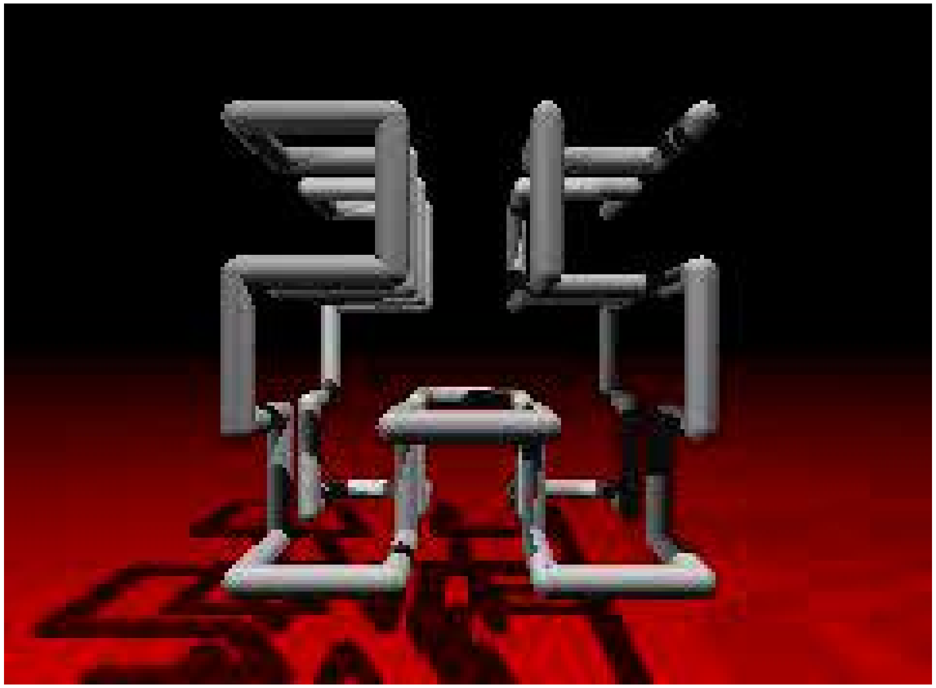
Recursive Inertial Bisection (Geometric)



Hilbert Space Filling Curve (Geometric)

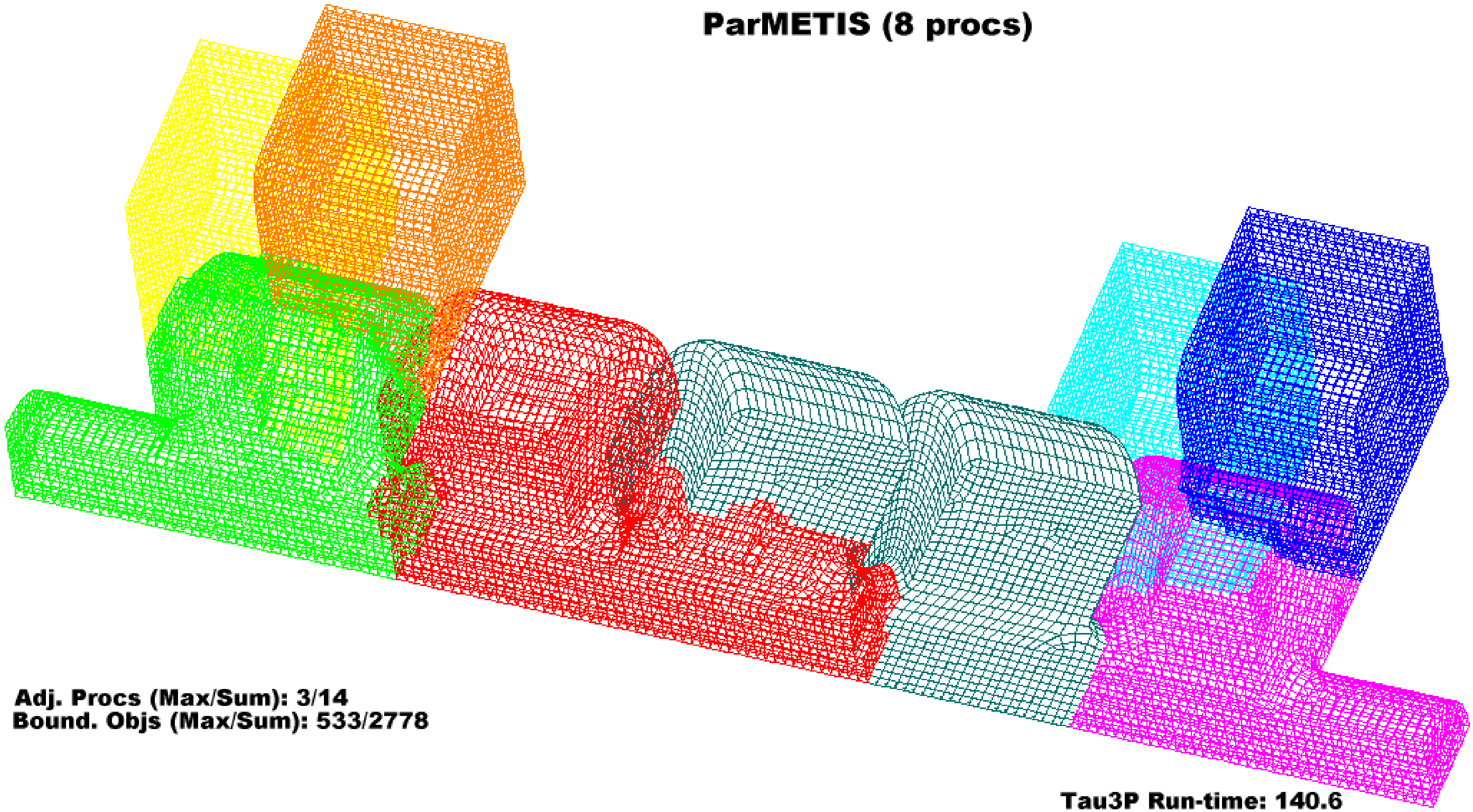


Hilbert Space Filling Curve (Geometric)

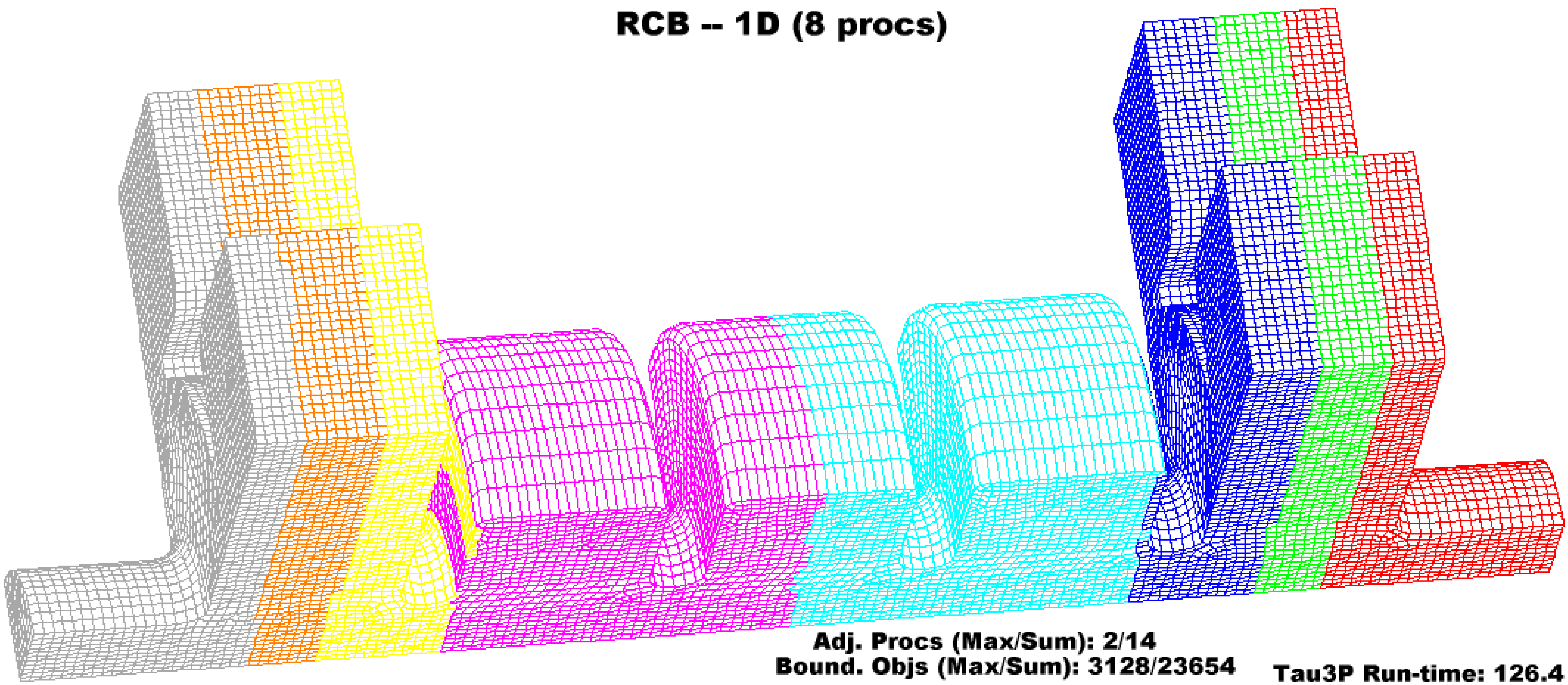


Tau3P Partitioning Results

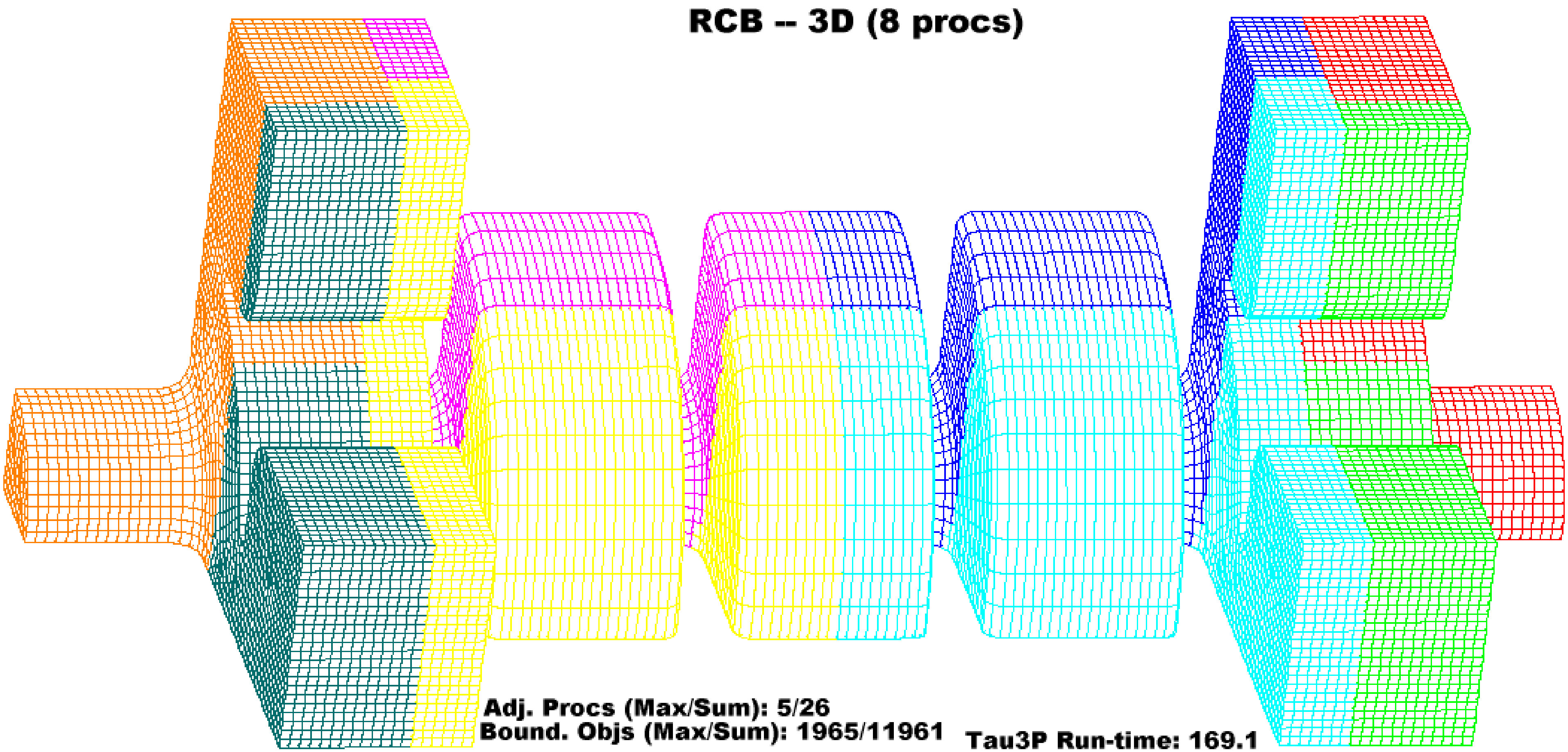
RDDS (5 cell w/ couplers) ParMETIS Partition



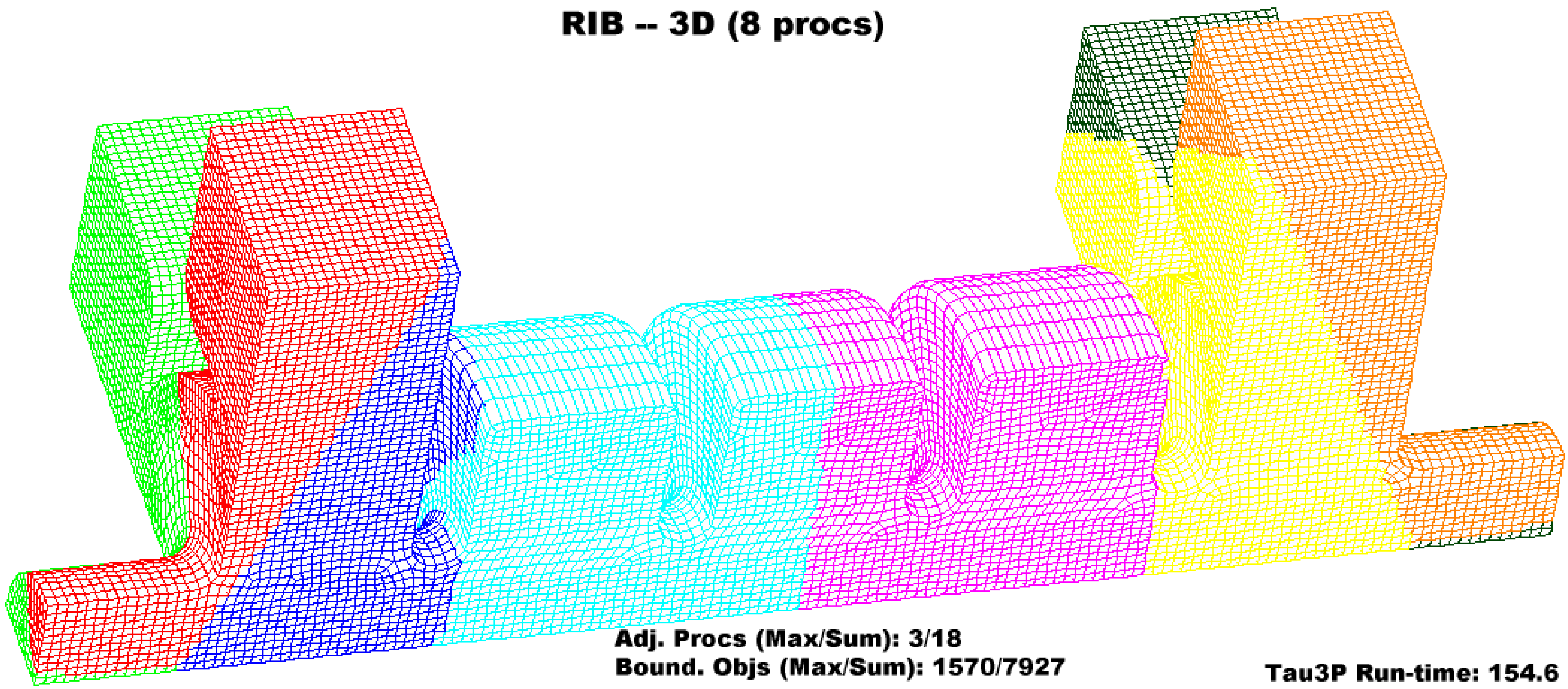
RDDS (5 cell w/ couplers) RCB-1D(z) Partition



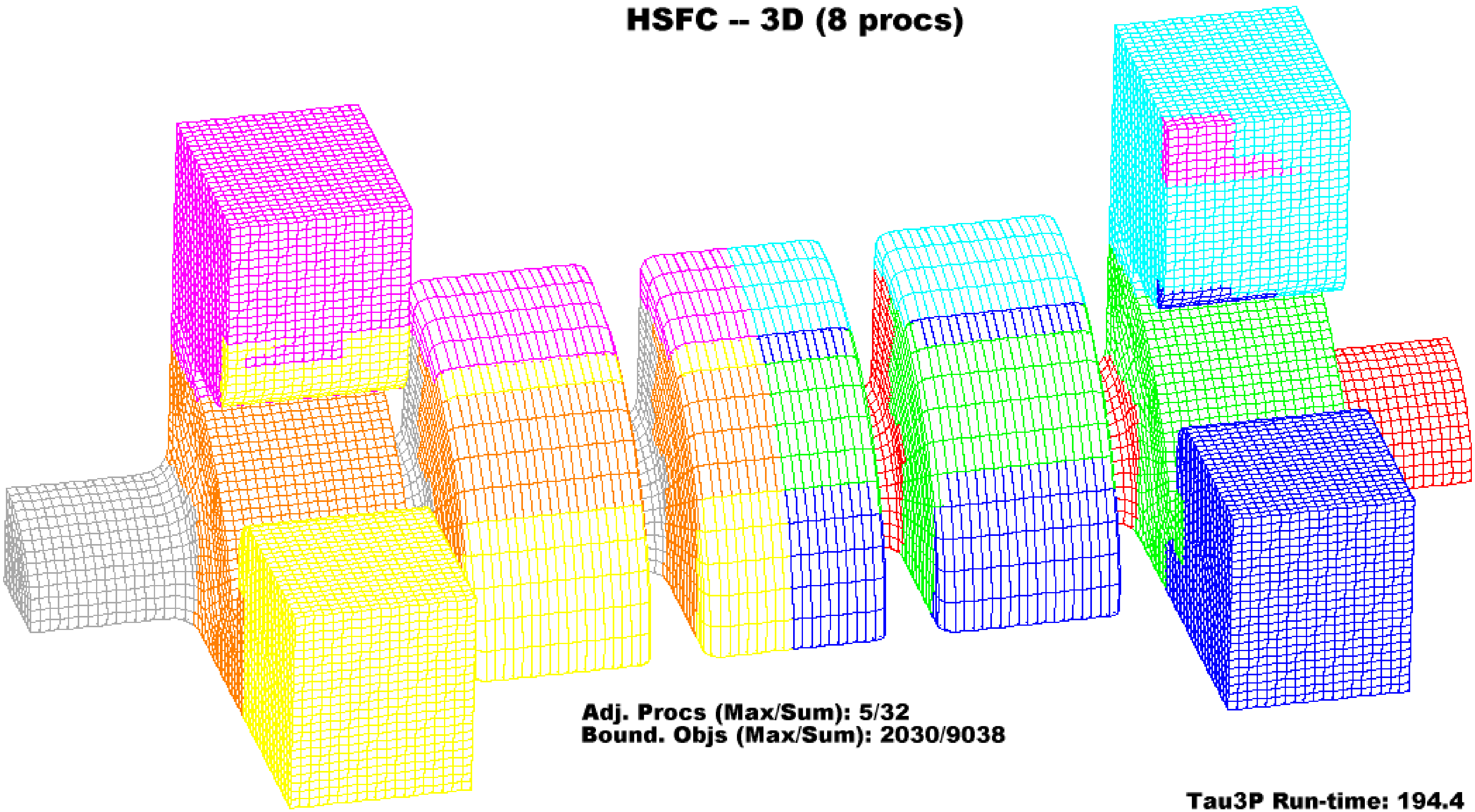
RDDS (5 cell w/ couplers) RCB-3D Partition



RDDS (5 cell w/ couplers) RIB-3D Partition



RDDS (5 cell w/ couplers) HSFC-3D Partition



5 Cell RDDS (8 processors) Partitioning

	Tau3P Runtime	Max Adj. Procs	Sum Adj. Procs	Max Bound. Objs	Sum Bound. Objs
ParMETIS	288.5 s	3	14	585	2909
RCB-1D (z)	218.5 s	2	14	3128	14363
RCB-3D	343.0 s	5	26	1965	11961
RIB-3D	282.4 s	3	18	1570	7927
HSFC-3D	387.3 s	5	32	2030	9038

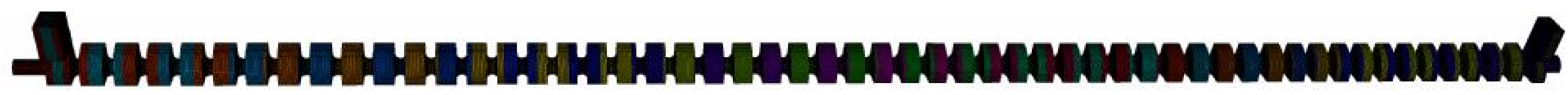
2.0 ns runtime
IBM SP3 (NERSC)

5 Cell RDDS (32 processors) Partitioning

	Tau3P Runtime	Max Adj. Procs	Sum Adj. Procs	Max Bound. Objs	Sum Bound. Objs
ParMETIS	165.5 s	8	134	731	16405
RCB-1D (z)	67.7 s	3	66	2683	63510
RCB-3D	373.2 s	10	208	1404	24321
RIB-3D	266.8 s	8	162	808	20156
HSFC-3D	272.2 s	10	202	1279	26684

2.0 ns runtime
IBM SP3 (NERSC)

H60VG3 ("real" structure)



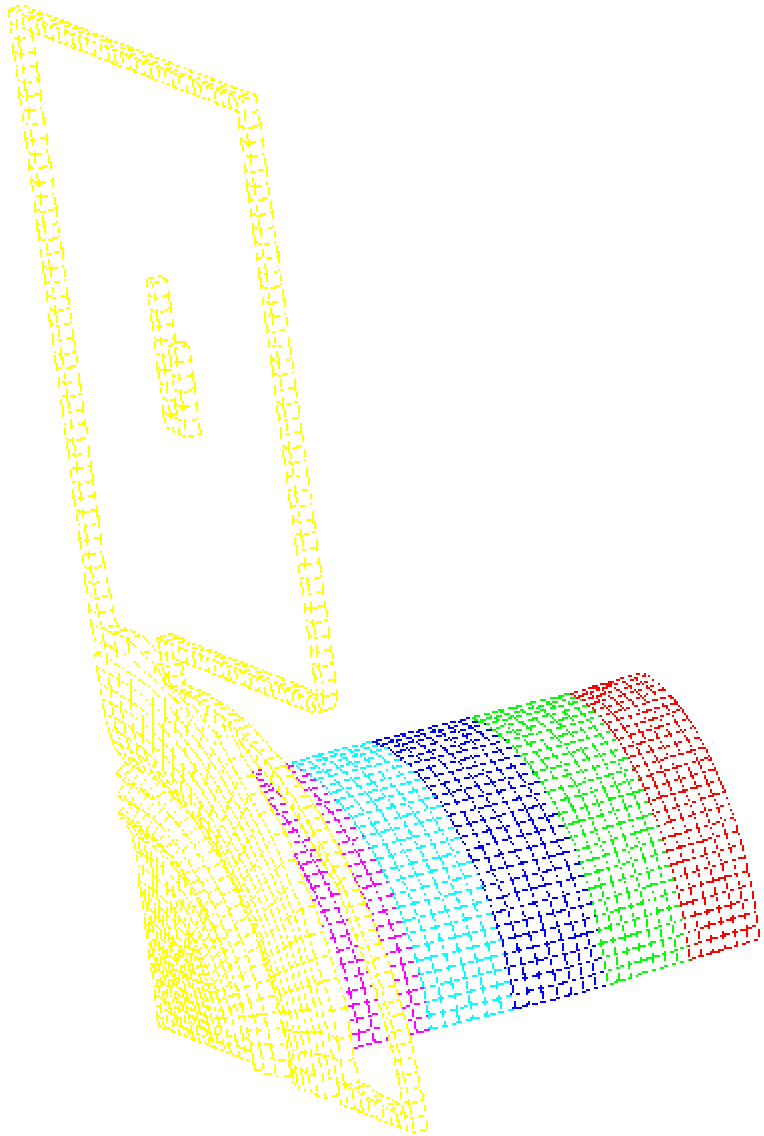
55 cells (w/ coupler)
1,122,445 elements

H60VG3 RDDS Partitioning (w/o port grouping)

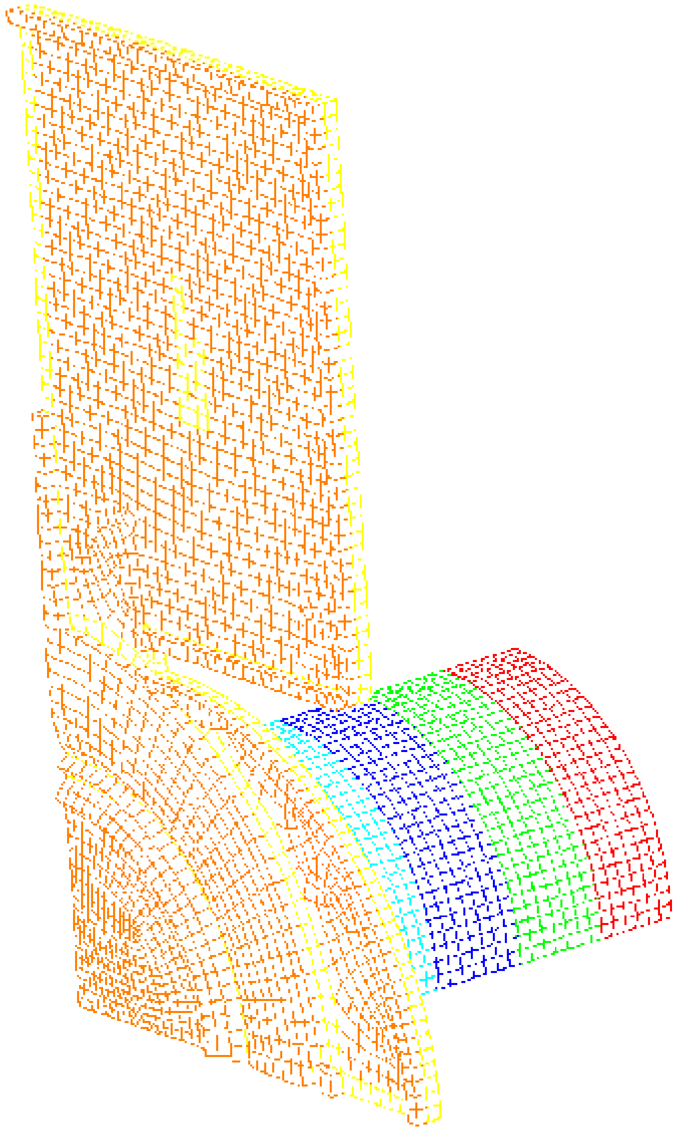
# of Procs	ParMETIS Max Adj. Procs	ParMETIS Speedup	ParMETIS Runtime	RCB-1D Runtime	RCB-1D Speedup	RCB-1D Max Adj. Procs
8	2	8/8	3930.7 s	3898.6 s	8/8	2
16	3	11.6/16	2703.3 s	2458.5 s	12.7/16	2
32	4	21.6/32	1455.0 s	1236.6 s	25.2/32	2
64	4	42.7/64	736.6 s	627.2 s	49.7/64	2
128	10	48.9/128	643.0 s	265.1 s	117.6/128	2
256	11	87.3/256	360.0 s	129.2 s	241.4/256	2
512	14	107.7/512	292.1 s	92.3 s	337.9/512	4
1024	16	96.0/1024	327.5 s	99.0 s	315.0/1024	8

1.0 ns runtime
IBM SP3 (NERSC)

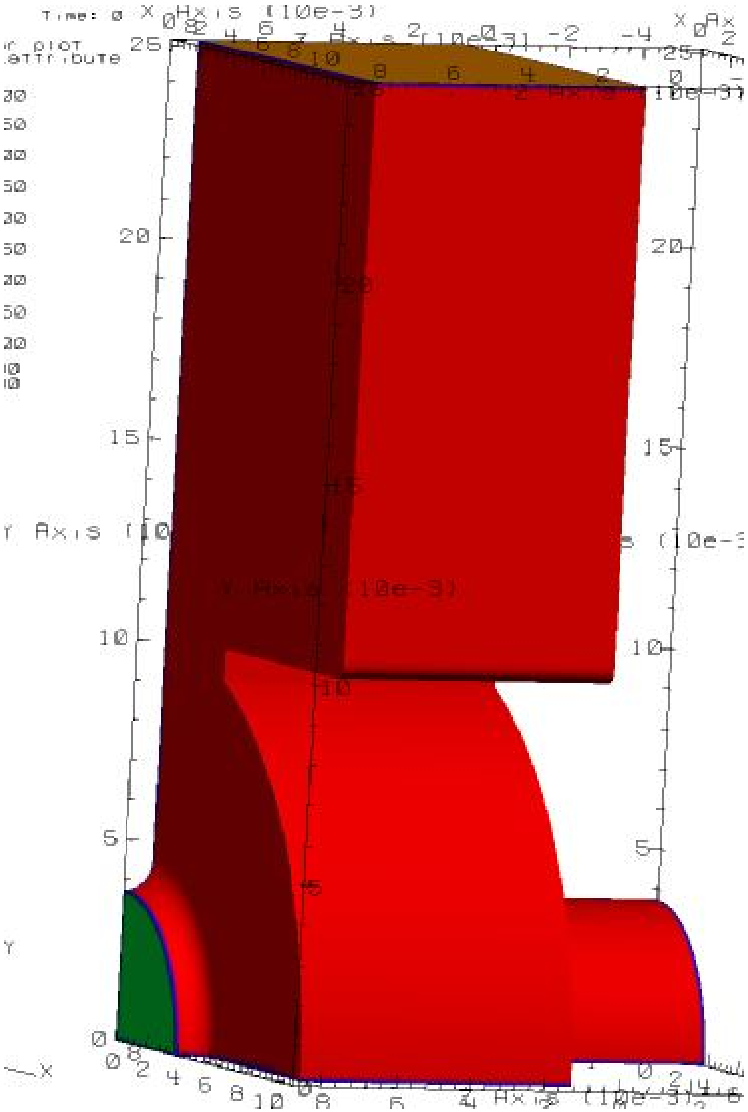
RCB Scalability Leveling Off



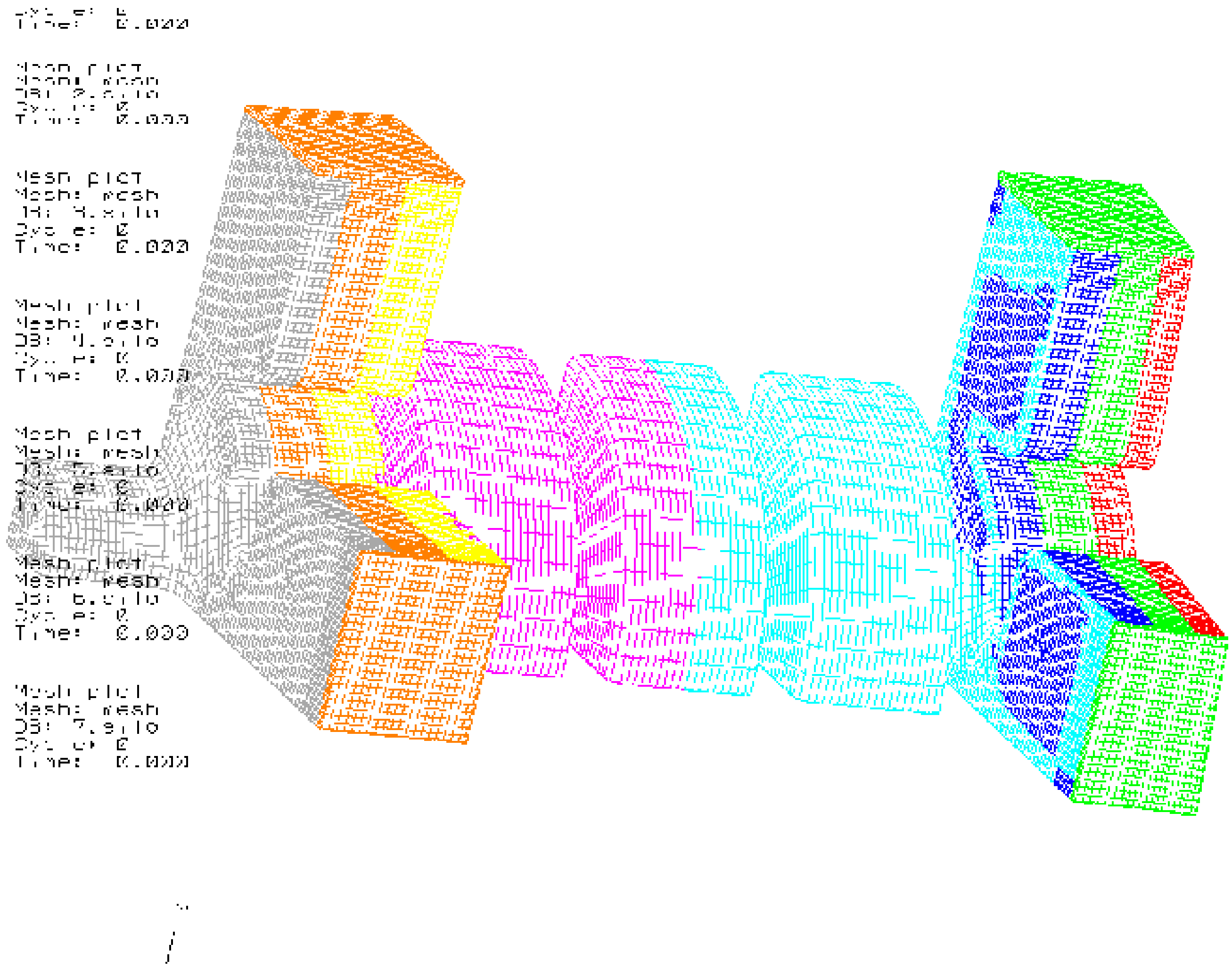
RCB Scalability Leveling Off



Coupler Port Grouping Complication



Coupler Port Grouping Complication

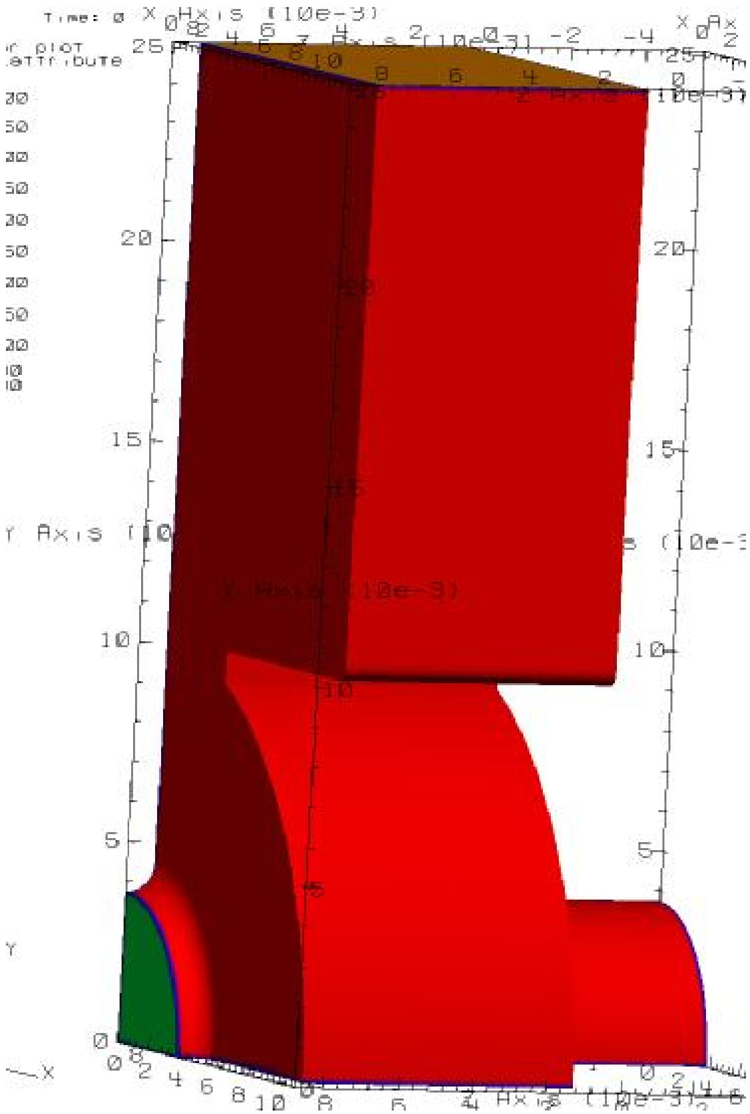


H60VG3 RDDS Partitioning (w/ coupler port grouping)

# of Procs	ParMETIS Max Adj. Procs	ParMETIS Speedup	ParMETIS Runtime	RCB-1D Runtime	RCB-1D Speedup	RCB-1D Max Adj. Procs
8	2	8/8	3856.2 s	3826.5 s	8/8	2
16	3	7.2/16	4257.0 s	2405.4 s	12.7/16	2
32	3	14.3/32	2158.3 s	1820.7 s	16.8/32	2
64	7	31.0/64	995.1 s	889.3 s	34.4/64	3
128	7	47.5/128	649.6 s	599.0 s	51.1/128	6
256	9	69.7/256	442.9 s	516.5 s	59.3/256	11
512	12	70.4/256	438.1 s	531.1 s	57.6/512	21

1.0 ns runtime
IBM SP3 (NERSC)

Constrained Mesh Partitioning



RDDS Coupler Cell Constrained Partition (16 procs)

Method	Max Adj. Procs
HSFC-3D	14
ParMETIS	8
RCB-1D-z	14
RCB-2D-xy	5
RCB-2D-xz	14
RCB-2D-yz	6
RCB-3D	8
RIB-2D-xy	6
RIB-2D-xz	14
RIB-2D-yz	5
RIB-3D	7

RDDS Coupler Cell Constrained Partition (32 procs)

Method	Max Adj. Procs
HSFC-3D	17
ParMETIS	14
RCB-1D-z	29
RCB-2D-xy	7
RCB-2D-xz	29
RCB-2D-yz	7
RCB-3D	11
RIB-2D-xy	7
RIB-2D-xz	29
RIB-2D-yz	6
RIB-3D	12

Future Work I Would Have Done

- Stitching Multiple Partitions Together
- Competition
- Onion Partition growing
- Dynamic Partitioning for Track3P

Acknowledgements

- SNL
  - Karen Devine, et al.
- LBNL
  - Ali Pinar
- SLAC (ACD)
  - Adam Guetz, Cho Ng, Lixin Ge, Greg Schussman, Kwok Ko

- Work supported by U.S. DOE ASCR & HENP Divisions under contract DE-AC0376SF00515