HIGH PERFORMANCE COMPUTING IN ACCELERATING STRUCTURE DESIGN AND ANALYSIS*

Zenghai Li†, Nathan Folwell, Lixin Ge, Adam Guetz, Valentin Ivanov, Marc Kowalski, Lie-Quan Lee, Cho-Kuen Ng, Greg Schussman, Lukas Stingelin[1], Ravindra Uplenchwar, Michael Wolf[2], Liling Xiao and Kwok Ko

Stanford Linear Accelerator Center, Stanford University, USA
[1]Paul Scherer Institute, Switzerland
[2]University of Illinois, Urbana-Champaign, USA

Abstract
Future high-energy accelerators such as the Next Linear Collider (NLC) will accelerate multi-bunch beams of high current and low emittance to obtain high luminosity, which put stringent requirements on the accelerating structures for efficiency and beam stability. While numerical modeling has been quite standard in accelerator R&D, designing the NLC accelerating structure required a new simulation capability because of the geometric complexity and level of accuracy involved. Under the US DOE Advanced Computing initiatives (first the Grand Challenge and now SciDAC), SLAC has developed a suite of electromagnetic codes based on unstructured grids and utilizing high performance computing to provide an advanced tool for modeling structures at accuracies and scales previously not possible. This paper will discuss the code development and computational science research (e.g. domain decomposition, scalable eigensolvers, adaptive mesh refinement) that have enabled the large-scale simulations needed for meeting the computational challenges posed by the NLC as well as projects such as the PEP-II and RIA. Numerical results will be presented to show how high performance computing has made a qualitative improvement in accelerator structure modeling for these accelerators, either at the component level (single cell optimization), or on the scale of an entire structure (beam heating and long range wakefields).

1. INTRODUCTION

Particle accelerators are among the most important and most complex scientific instruments in use, and are critical to research in fields such as high-energy physics, nuclear physics, materials science, chemistry, and the biosciences. As new and existing facilities continually strive towards higher energy, higher beam current, and greater efficiency, accelerator physicists and engineers are faced with increasingly demanding specifications on the RF system to improve performance and reduce cost. As a result, the emphasis of designing the accelerating structures for these machines has been placed heavily on numerical modeling as the cost-saving approach to their R&D. This means the accuracy and reliability of the modeling software are becoming of paramount importance in order that the structures can meet the stringent design requirements.

The Damped, Detuned Structure (DDS) shown in Fig. 1 is the baseline linac design for the warm Linear Collider (NLC)[1] scheme. In the DDS, the frequency of the accelerating field must be accurate to within 1 part in 10,000 to maintain acceleration efficiency. This requirement has to be met in a complex cavity geometry that optimizes the accelerating field gradient while suppresses the long-range dipole wakefields. To provide the desired accuracy in the DDS cell frequency and to verify the wakefield suppression by damping and detuning over the entire structure are modeling challenges that are beyond the capabilities of standard electromagnetic (EM) codes (e.g. MAFIA[2] and HFSS[3]) running on limited computing resources like desktop computers.

Figure 1. Model of the 55-cell NLC DDS design.

* Work supported by the U.S. DOE Contract No. DE-AC02-76SF00515.
† E-mail address: lizh@slac.stanford.edu

Contributedto8thInternationalComputationalAcceleratorPhysicsConference(ICAP2004),
06/29/2004--7/2/2004,St.Petersburg,Russia

WorksupportedinpartbyDepartmentofEnergycontractDE-AC02-76SF00515

2. HIGH PERFORMANCE COMPUTING

Besides the NLC DDS design, the PEP-II[4] Interaction Region (IR), the RIA[5] RFQ cavity and the PSI[6] ring cyclotron are additional examples for which more advanced simulation tools other than available software are needed. In the PEP-II, beam heating in the IR due to trapped modes is an obstacle to high current operation, and thus high luminosity. Fig. 2 shows the IR beamline complex which consists of a central vacuum chamber of complicated, varying cross sections (due to the synchrotron masks) that connects to the positron and electron beamlines via a crotch junction at both ends. Modeling the entire geometry is necessary to fully account for the beam heating effect. The proposed RIA plans to employ RFQ structures such as those shown in Fig. 3 in its low energy linacs. Presently, provisions have to be made for tuners to compensate for frequency errors of about 1% from using existing design software. New modeling tools that can improve the accuracy by an order of magnitude would lead to a significant reduction in tuners required and a much simplified operation as well.

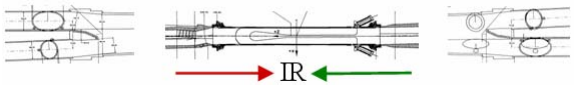

Figure 2. Beamline complex of the PEP-II Interaction Region (IR).

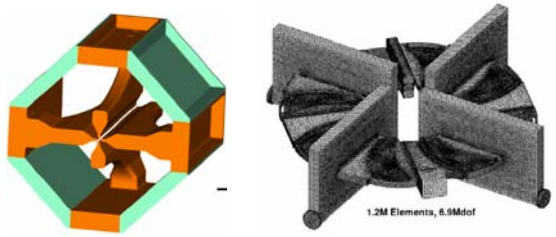

Figure 3. The RIA RFQ and the PSI ring cyclotron.

To address these modeling challenges, SLAC has embarked on a code development effort that was first initiated in 1997 under the support of the DOE Grand Challenge on Accelerator Physics, with the goal to implement high performance computing (HPC) capabilities in accelerating modeling tools. This work has expanded with the follow-up DOE SciDAC Accelerator Simulation project in which SLAC leads the team that specifically targets high accuracy, large-scale electromagnetic applications.

3. PARALELL ELECTROMAGNETIC CODES ON UNSTRUCTURED GRIDS

The suite of 3D, parallel electromagnetic codes that are finite element based consists of:

(1) **Omega3P** – eigenmode solver for finding normal modes in lossless and lossy cavities,

(2) **S3P** - solver in frequency domain to calculate S parameters of open structures,

(3) **T3P** – time-domain solver for modeling response due to beam, dipole and waveguide excitation,

and it also includes:

(4) **Tau3P** – time-domain solver following the Discrete Surface Integral (DSI) formulation with same functionalities as **T3P,**

(5) **Track3P** – module for dark current simulation with surface physics using fields from solvers above,

(6) **Viz3D** – analysis and graphics package.

The finite element codes employ tetrahedral mesh elements while the DSI based **Tau3P** uses hexahedral cells (Fig. 4), both unstructured grids able to conform to curved surfaces for very high accuracy modeling. . The set of codes is developed under C++ with a unified data structure to facilitate geometry input and partitioning, and uses MPI for communication on distributed memory architectures.

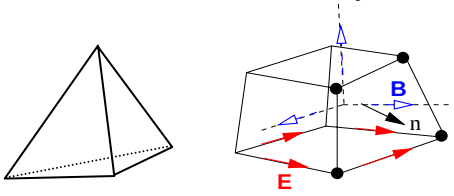

Figure 4. Tetrahedral mesh for Omega3P/S3P/T3P and hexahedral mesh for Tau3P.

The code development at SLAC is supported by an extensive and coordinated R&D program in computer science and applied mathematics that is sponsored by SciDAC and carried out in designated national laboratories and universities. Among these efforts include parallel meshing at Sandia National Lab (SNL) and U. of Wisconsin, partitioning at Sandia and Lawrence Berkeley Lab (LBL), linear solvers and eigensolvers at Stanford and LBL, adaptive refinement at RPI, and visualization at UC Davis. They contribute to the success of the large-scale simulations required for the challenging accelerator applications described above. We will next present two examples of high resolution

modeling and three examples of system scale simulations using the new HPC tools.

4. HIGH RESOLUTION DESIGN

The DDS cell is one of 206 cavities in the NLC first accelerating structure design whose cavity shape is optimized for high shunt impedance while connected to four damping manifolds that run the length of the structure for long-range dipole wakefield suppression. The wake is further reduced by a cavity-to-cavity variation that detunes the dipole modes in a Gaussian manner. As previously indicated, the fundamental mode frequency of the DDS cell has to be accurate to 0.01% to maintain acceleration efficiency. Applying **Omega3P** to the distributed model shown in Fig. 5 on NERSC's IBM/SP, a table of dimensions for all 206 different cavities along the DDS was generated for computerized machining based on calculations that met this accuracy requirement. Cold tests on fabricated cells based on this table indeed showed that their measured frequencies are within the targeted value (Fig. 5). This result demonstrates that high-resolution modeling utilizing unstructured grids and parallel computing is a powerful tool for designing the NLC structures which potentially could have saved the project more than $100 million in machine cost alone. Furthermore, the resolution and speedup provided by HPC together with high precision machining have enabled simulation based design to become a cheaper, faster alternative to the traditional expensive and time-consuming R&D process of repeated fabrication and testing.

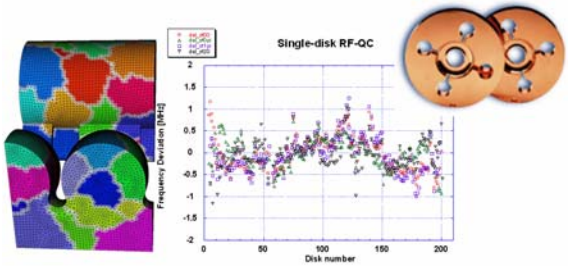

Figure 5. (Left) A distributed Omega3P model of 1/8 of the DDS cell, (right) microwave QC of the fabricated cells (insert) showing measured frequencies within 0.01% of target value.

Under SciDAC, SLAC and RPI are collaborating on developing an adaptive mesh refinement (AMR) capability in **Omega3P** to improve the accuracy and convergence of frequency and quality factor calculations for cavities of complex shapes. Accurate wall loss determinations are difficult when the wall currents are localized in narrow regions of the cavity, such as around the coupling iris. The increased wall loss reduces the cavity's quality factor, thus the shunt impedance, and can also lead to RF surface heating at high power. An adaptive mesh control loop based on error indication procedures has been implemented in **Omega3P** to provide increasingly refined meshes until a converged result is obtained. The adaptive procedure has been applied to the RIA RFQ cavity with the following results. Using measurement[7] as benchmark, the percentage deviations in frequency and Q from a standard code are 1.5 and 30 respectively while the results from **Omega3P** using uniform refinement are 0.05 and 16.1 and with AMR the numbers are 0.11 and 16.2. The AMR case uses 0.4 million DOFs, 10 times less than the uniform refinement case to reach the same accuracy. We conclude that combining unstructured grid, 3rd order elements and AMR**, Omega3P**'s improvements over standard codes are an order of magnitude better in frequency and about a factor of two better in Q.

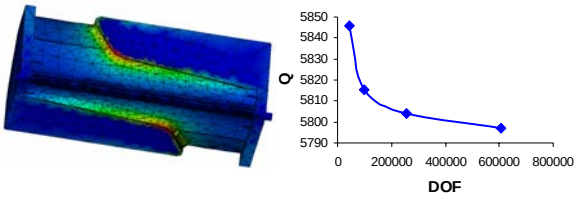

Figure 6. Adaptively refined mesh and wall loss for the RFQ cavity and the convergence of Q vs DOF.

5. SYSTEM SCALE MODELING (DDS)

The h60vg3[8] X-band structure is a 55-cell DDS that is considered to be the baseline design for the NLC linac. Due to the complicated geometry, the long-range dipole wakefields in this structure have thus far been analyzed by an approximate equivalent circuit model representing only the lowest two dipole bands. It is of both theoretical and practical interests to be able to simulate the entire, realistic structure so that the effect of the higher bands and that due to the input/output couplers can be included. Modeling a problem of such size and complexity has not been tried before because the computational resources required were considered prohibitive. Using the parallel tools developed under SciDAC and the computing resources it provides on

NERSC's IBM SP2 machine, such a system-scale simulation is within reach. We will next present the computations of the wakefields in the h60vg3 structure via time and frequency domain methods.

Time-domain Simulation: The parallel time-domain solver **Tau3P** has been used to simulate the h60vg3 structure with a transit beam so that the wakefield can be found directly for the first time. Fig. 7 shows the beam-excited electric fields in the structure at two instances in time. The top snapshot shows the field excitation as the beam traverses the structure. The bottom one is the field distribution after the beam left the structure, clearly showing the coupling of the wakefield out to the manifold and subsequently out of the structure through the HOM coupler at the downstream end. The evolution of the wakefield from the **Tau3P** simulation provides insight into the HOM damping via the manifold.

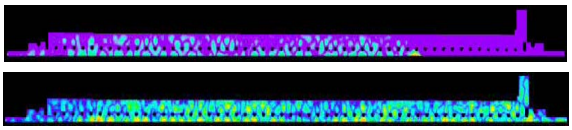

Figure 7. Time domain **Tau3P** simulation of a rigid beam traversing the DDS structure, showing field distribution at two instances in time.

The wakefield is plotted in Fig. 8 together with the results from the same structure but without the manifolds. The difference between the DS (Detuned Structure) and DDS wakefields due to manifold damping at long distances is clearly evident. Fig. 9 is the impedance spectrum showing the contributions from the higher bands up to 30 GHz which are absent from the two-band equivalent circuit model.

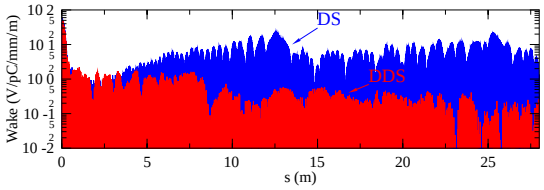

Figure 8. Dipole wakefield in the 55-cell h60vg3 DDS in red as compared with that of the same structure but without the manifold in blue.

Frequency-domain Calculation: The wakefield in the h60vg3 structure can also be found by summing

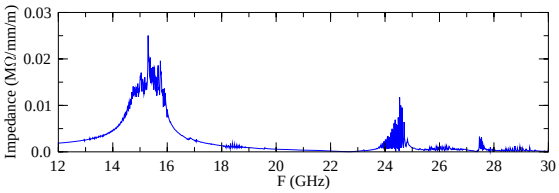

Figure 9. Impedance spectrum of h60vg3 wakefield from **Tau3P** simulation covering up to 30 GHz.

the eigenmodes in the frequency domain. Because of the power loss through the fundamental and HOM couplers, a complex solver has been developed in **Omega3P** to calculate damped modes. In the computational model lossy materials are used as matched terminations for the fundamental and HOM couplers in the frequency range of interest. About 400 complex eigenmodes covering the spectrum up to 30 GHz. have been calculated to form the sum for the wakefield. In Fig. 10 are shown two typical modes in the DDS structure. The top plot is a mode that couples to the backward wave in the manifold, suggesting the need for a HOM load at the input end. The bottom plot shows a mode that couples to a forward wave in the manifold that is damped by the HOM load downstream. This is the first time these modes have been seen in a long, realistic structure.

The mode spectrum from **Omega3P as** depicted in Fig. 11 shows good agreement when compared with the **Tau3P** result of Fig. 9. The agreement is even closer on the wakefields as shown in Fig. 12 which is remarkable considering the results have been obtained from two different methods each with its mesh model of the structure. This serves not only to validate the two codes but to also verify the effectiveness of the detuning and damping scheme for wakefield suppression. Furthermore, it points out the importance of high performance computing and the SciDAC support in computational science research, without which such calculations would not have been possible.

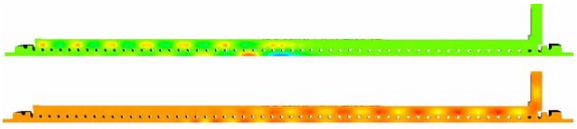

Figure 10. (Top) Complex eigenmode in the h60vg3 structure that couples to a backward wave in the manifold; (bottom) mode couples to a forward wave that gets damped via the HOM coupler down stream.

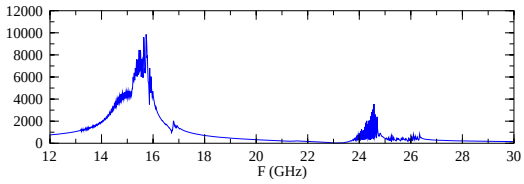

Figure 11. Impedance spectrum of h60vg3 wakefield from **Omega3P** analysis.

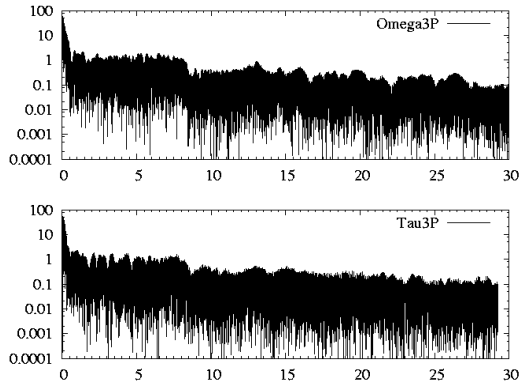

Figure 12. **Omega3P** and **Tau3P** wakefields out to 30 m behind leading bunch, showing detuning effect at shot range, and damping effect at long range.

Advances in Eigenmode Solvers: The success of the large-scale simulations just presented relies heavily on the accuracy, convergence and scalability of the solvers used. SciDAC supports a strong effort consisting of SLAC, Stanford and LBL to develop new, efficient algorithms for solving the most challenging applications. They include a linear solver framework that interfaces **Omega3P/S3P** to direct solvers such as SuperLU and WSMP, leading to an improvement in speed by a factor of 50-100. Implementation of the AV FEM formulation adds a gain of 5-10 for using iterative solvers. In addition, higher-order hierarchical bases (up to p=6) and preconditioners have been developed to increase accuracy and convergence. With these advances we have been able to solve the h60vg3 structure using up to 93 million DOFs, requiring about 800 GB of memory (Fig. 13 left), and they helped in reaching higher accuracy in AMR calculations. Furthermore, work on complex solvers has enabled the mode analysis of the h60vg3 and the PEP-II IR absorber design to be discussed later.

Research in Domain Decomposition: The scalability of parallel time-domain codes such as **Tau3P** is hampered by communication costs since the fields have to be updated at every time step, leading to poor parallel efficiency. Under SciDAC SLAC is working with SNL and LBL to improve **Tau3P** performance by exploring alternate schemes within the Zoltan partitioning library. For the **Tau3P** simulation of the h60vg3 55-cell structure, Fig, 13 (right) shows that the RCB1D partitioning provides better speedup over the existing partitioning using ParMETIS by reducing the number of neighboring MPI processes.

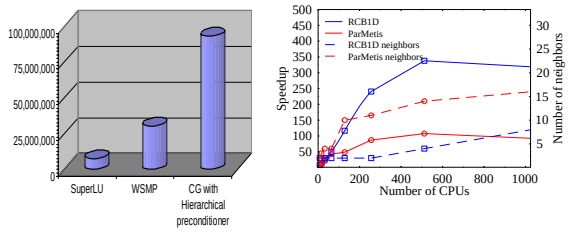

Figure 13. (Left) Largest problem size in DOFs with three linear solvers. Maximum reached is 93 million; (right) Speedup gain using RCB1D over ParMETIS.

6. SIMULATING THE PEP-II IR

A distributed model of the PEP-II IR shown in Fig. 14 is used in **Tau3P** simulation of two colliding beams traversing the beamline complex near the interaction point. Two snapshots of the simulation before and after the beams pass each other are shown in Fig. 15. The beam-excited fields are analyzed to find the trapped modes that contribute to local heating effects. These heating calculation were factored into the upgrade of the IR and the machine now is able to operate at 15% higher beam current without overheating.

Figure 14. Distributed model of PEP-II IR for **Tau3P** simulation of beam heating.

Figure 15. Beam-excited fields in PEP-II IR from **Tau3P** simulation before and after two beams passing each other at the collision point.

In preparation for further increase in beam current a new IR design is under consideration that uses absorbers to damp trapped modes and reduce beam heating. The absorbers are located at the

crotch region on both ends of the central beam pipe (see Fig. 16). Their effectiveness is studied using the complex eigensolver in **Omega3P.** Fig. 17 compares the quality factor of the trapped modes between the crotches with and without the absorbers. It is seen that most of the modes have been damped by two orders of magnitude. This simulation capability provides a powerful tool for optimizing the absorber design without extensive prototyping and testing.

Figure 16. **Omega3P** model of the PEP-II IR including placement of absorbers at the crotches.

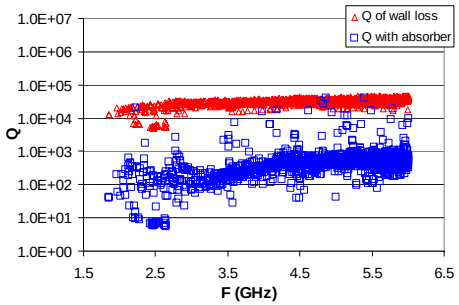

Figure 17. Complex **Omega3P** calculation of damping effect on trapped modes in PEP-II IR using absorbers at the crotches (undamped modes in red; damped modes in blue)

7. MODELING ENTIRE RING CYCLOTRON

The availability of new parallel codes such as **Omega3P** has generated interests in finding the HOM modes that exist in an entire ring cyclotron for better understanding of their effect on the beam dynamics in such a machine. PSI and SLAC are collaborating on modeling the PSI ring cyclotron with **Omega3P** and determining the HOM effects through an eigenmode analysis. The calculations have been performed using the 32 CPU IBM/SP4 at PSI. Using the ESIL solver in **Omega3P**, 280 modes with frequency close to a beam harmonic have been computed. They can be classified into three types. Forty four of them are cavity modes with low frequency and high gap voltage. There are eighteen vacuum chamber modes having medium frequency and low gap voltage. The rest of the modes are hybrid modes that occupy both the cavity and the vacuum chamber, and these have high frequency and low gap voltage. In Fig 18, sample field pattern of the three types of modes in an entire ring cyclotron is shown for the first time.

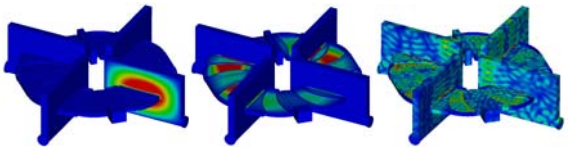

Figure 18. Eigenmodes in the PSI Ring Cyclotron: cavity mode (left), vacuum chamber mode (middle), and hybrid mode (left).

SUMMARY

A new suite of parallel electromagnetic codes based on unstructured grids has been developed under the DOE SciDAC project to model large, complex accelerator cavities and RF structures that have not been possible with standard software. It is demonstrated that the high performance computing capability in these codes has enabled the large-scale simulations that are necessary for high resolution cavity design and system-scale structure analysis. R&D in computer science and applied mathematics through SciDAC collaborations have shown to further improve their accuracy and efficiency. The codes have been applied to optimize existing accelerators and design planned facilities with considerable success.

REFERENCES

[1] ZDR report for the Next Linear Collider, LBNL-PUB-5424, SLAC Report 474, UCRL-ID-124161, 1996.

[2] MAFIA User's Guide, CST GmbH, Darmstadt, Germany.

[3] HFSS: http://www.ansoft.com/products/hf/hfss

[4] PEP-II: http://www.slac.stanford.edu/accel/pepii

[5] RIA: http://www.orau.org/ria

[6] PSI: http://gfa.web.psi.ch

[7] N. Vinogradov, et al, Progress with the room temperature structures for the RIA facility, RIA R&D Workshop, August 26-28, 2003.

[8] Z. Li, et al, Optimization of the x-band structure for the JLC/NLC, proceedings of PAC2003, Portland, 2003.