Y3 Annual Report
1/27/2016-1/26/2017
# Table of Contents

1. Introduction ............................................................................................................................. 4  
   1.1 Background ...................................................................................................................... 4  
   1.2 Demonstration Problem ............................................................................................... 6  
   1.3 Simulation Roadmap ....................................................................................................... 7  
   1.4 Integration ..................................................................................................................... 9  
2. Macroscale Team ................................................................................................................... 13  
   2.1 Overview ........................................................................................................................ 13  
   2.2 Progress on the Demonstration Problem ....................................................................... 13  
   2.3 Code Improvement ........................................................................................................ 14  
   2.4 Collaborative Effort ....................................................................................................... 14  
   2.5 Fourth Year Plans ........................................................................................................... 15  
3. Microscale Team ................................................................................................................... 19  
   3.1 Goals and Motivation .................................................................................................... 19  
   3.2 Shock Interaction with Structured Arrays of Particles ................................................... 19  
      3.2.1 Shock interaction with single particle ..................................................................... 19  
      3.2.2 Shock interaction with one dimensional structured array of particles ................. 21  
      3.2.3 Shock interaction with one layer of transverse array of nine particles ............... 21  
      3.2.4 Shock interaction with FCC array of particles ...................................................... 22  
   3.3 Shock Interaction with Random pack of particles ........................................................... 26  
   3.4 Interaction of Contact Discontinuities with Particles .................................................... 30  
   3.5 Fourth Year Plans ........................................................................................................... 32  
4. Experiments ........................................................................................................................... 33  
   4.1 ASU Experiments ........................................................................................................... 33  
      4.1.1 Goals and Motivation ............................................................................................. 33  
      4.1.2 General Process ....................................................................................................... 33  
      4.1.3 Particle Image Velocimetry .................................................................................... 34  
      4.1.4 Air-particle bed Interface Tracking ....................................................................... 35  
      4.1.5 Cellular Pattern ....................................................................................................... 36  
      4.1.6 Cloud Formation .................................................................................................... 37  
      4.1.7 Equipment Purchases ............................................................................................. 39  
      4.1.8 Fourth Year Plans .................................................................................................... 40  
   4.2 Eglin AFB Experiments .................................................................................................. 41
1. Introduction

1.1 Background

The University of Florida (UF) established a Center for Compressible Multiphase Turbulence (CCMT) on January 26, 2014 as part of the NNSA's Predictive Science Academic Alliance Program II (PSAAP-II) Single-Discipline Centers (SDC). The intellectual objectives of the Center are threefold: to radically advance the field of compressible multiphase turbulence (CMT) through rigorous first-principle multiscale modeling, to advance very large-scale predictive simulation science on present and near-future platforms, and to advance a co-design strategy that combines exascale emulation with a novel energy-constrained numerical approach. The Center is performing petascale, and working towards exascale, simulations of instabilities, turbulence and mixing in particulate-laden flows under conditions of extreme pressure and temperature to investigate fundamental problems of interest to national technological leadership. Towards this vision we are tackling the following challenges:

1) Target an important application that can only be enabled by exascale computing: We are solving a complex multiscale problem at an unprecedented level of physical detail and integration and thereby advance predictive simulation science. CMT poses a grand challenge to our understanding as it combines three complex physics: compressibility, multiphase flow and turbulence. CMT occurs often under extreme conditions of pressure and temperature, and as a result is not easily amenable to high-fidelity experiments and diagnostics. CMT presents a fascinating array of poorly-understood instability, transition, and turbulent processes manifest over a wide range of strongly interacting length and time scales. Current computational approaches involve models and closures that are developed from incomplete understanding, and as a result are largely empirical. Fully validated exascale simulation perhaps is the only path to fundamental breakthroughs that can lead us out of current empiricism.

2) Well-defined problem hierarchy leading to a demonstration problem: A multiscale approach from the microscale to the mesoscale and to the macroscale is being pursued for a systematic integrated investigation of the CMT physics. We have adopted a problem hierarchy that culminates at a signature demonstration problem of explosive dispersal of particles from a well-characterized initial condition, which fully exercises all the key complex processes of CMT. We pursue a coupling strategy where (i) fully resolved microscale simulations will lead to reduced order descriptions (interphase coupling models) to be employed at the mesoscale and (ii) partially resolved mesoscale simulations will lead to reduced order descriptions (multiphase large eddy

Goals of CCMT
- To radically advance the field of CMT
- To advance predictive simulation science on current and near-future computing platforms with uncertainty budget as backbone
- To advance a co-design strategy that combines exascale emulation, exascale algorithms, exascale CS
- To educate students and postdocs in exascale simulation science and place them at NNSA laboratories
simulation closures) to be employed at the macroscale. This will allow computational efficiency and high degree of parallelism at all levels of the hierarchy.

3) Simulation and experiment roadmaps for rigorous validation: We focus on integrated system-scale simulations of the demonstration problem from the outset using existing integrated code capabilities. Simultaneously, we also perform petascale simulations at the micro and mesoscales. Improvements to micro-to-meso and meso-to-macro coupling models will be systematically and periodically incorporated at the appropriate higher level. A layered systems engineering approach is used to organize and integrate physical subsystems with numerical, software and service components, to achieve progressively improved operational capability for system-scale simulations. We have developed a detailed simulation and experiment roadmap which allow rigorous step-by-step validation at each step of the problem hierarchy.

4) Develop novel uncertainty quantification (UQ) approaches for CMT: Detailed measurements from carefully chosen existing and planned experiments at the Air Force Research Laboratory Munitions Directorate (AFRL-RW), Sandia Multiphase Shock Tube facility and Los Alamos Center of Mixing under Extreme Conditions (CoMuEX) are used for rigorous quantification of uncertainties from the micro/mesoscales to the macroscale. We are engaged in vigorous uncertainty reduction through better characterization and instrumentation, rigorous calibration of the models, and improved numerical resolution. Simultaneous simulations and experiments at the micro, meso and macroscales of the problem hierarchy will allow us to both propagate up uncertainty to higher scales, and to reduce uncertainty through iterative improvements at the lower scales. A particularly difficult aspect of CMT is that it is characterized by extreme events that are localized in space and time. A key innovation is the development of novel techniques for accurate characterization of probability tails in the uncertainty quantification of such rare but critical events.

5) Demonstrate integrated performance on current/near-future architectures: Modern many-core architectures (such as Intel MIC), that provide high raw gigaflops, have deep memory hierarchies and low overhead threading capabilities. We exploit these capabilities to optimally utilize both computational and energy resources. In particular, we will tackle load balance and performance challenges in terms of data and work decomposition for the CMT code framework. Different parallelization schemes will be considered for effectively implementing simulations at the microscale, mesoscale, and system-scale, especially for heterogeneous resources.

6) Develop methods for predicting performance on a variety of exascale architectures: While many exascale trends seem clear, there are far too many permutations in the design space to select one a priori. We leverage the unique Novo-G facility at the NSF-supported UF Center for High-Performance Reconfigurable Computing (CHREC) to emulate and evaluate a series of candidate exascale architectures. We are developing an unprecedented capability to behaviorally prototype in software and hardware a variety of promising (as defined by leading exascale initiatives) forms of next-generation exascale (i) device and node designs at the micro-level and (ii) communication and system architectures at the macro-level. We are conducting experiments with CMT-bone
kernels, miniapps and skeleton-apps to evaluate promising architectures in terms of performance, energy, temperature, reliability, and scalability. Modeling, simulation, and estimation tools (e.g., those supported within the Sandia’s Structural Simulation Toolkit (SST)) are being leveraged with our behavioral simulations and emulations.

7) Solutions for energy efficiency and thermal management: We are developing a framework for multi-element and multi-objective optimization that will simultaneously minimize energy and maximize performance. We exploit the data and task parallelisms within CMT application and its UQ implementation to develop innovative low complexity static and dynamic algorithms for scheduling, while considering important factors such as thermal constraints and leakage currents.

1.2 Demonstration Problem

We aim at solving a problem of Compressible Multiphase Turbulence (CMT) at an unprecedented level of physical detail and thereby advance predictive simulation science. The overarching demonstration problem consists of a cylindrical core of simple explosive pellet of about 10 grams will be surrounded by a cylindrical very-thin-walled glass jacket of larger diameter. The annular region between the pellet and the jacket will be filled with mono or polydisperse metal powder of spherical shape. The shape and amount of the explosive charge and the size distribution of the metal powder and its material (aluminum, steel, tungsten, etc.) are parameters that will be varied. The charge will be hung from a test fixture so that the effect of the ground and the surrounding structures will be eliminated during the initial phase of explosion and dispersion. The orientation of the test setup will be such that the resulting explosive dispersal of particles and the gas field can be highly accurately measured. The following features makes this problem a very good choice for demonstration: (i) the explosive dispersal exercises all the major CMT physics, (ii) the extreme conditions makes this a demanding test for predictive capability, (iii) this problem requires exascale for true predictive capability, and (iv) we have already performed similar experiments

Figure 1.1. Left: Schematic of the three-dimensional computational domain for the explosive dispersal of particles simulations. Not drawn to scale. Right: Current three-dimensional hero run: 30 million computational cells, 5 million computational particles, computational domain extended to about 4m in radius. Time shown, t=1.46ms.
and validation-quality measurements. The explosive dispersal of solid particles problem displayed in Figure 1.1 and described by Frost et al. (Phys. Fluids, 24(9), 2012) was chosen for the initial phase of our research activities.

1.3 Simulation Roadmap

The center is focused on integrated system-scale simulations of the demonstration problem from the outset using existing integrated-code capabilities. Figure 1.2 shows the roadmap of the proposed sequence of simulations. The following important considerations was used in constructing the roadmap: (i) Along with system-level simulations of the demonstration problem, we will perform increasingly more complex simulations at the micro and mesoscales. Based on these simulations, improvements will be made to micro-to-meso and meso-to-macro coupling

![Roadmap Diagram]

Figure 1.2. A roadmap showing the sequence of progressively more complex simulations to be performed over the five years of the center. Micro and mesoscale simulations will be performed along with macroscale simulations of the demonstration problem. Both hero runs and bundled runs will be performed. Bundled runs will be combinations of varying parameters and ensemble of realizations. This plan is based on future petascale platforms that are expected to be available through NNSA. Concurrent simulations will be accomplished using DAKOTA.
models. (ii) To take maximum advantage of validation experiments, large numbers of simulations will be required for optimal calibration. We are using surrogate models to allow us to solve the multi-level optimization problem associated with selecting the physical constants that give the best match with the numerical model. (iii) Variations of the key control parameters (particle size, particle material, shock strength, etc.) will be guided by simulations that identify which combinations of parameters will elicit different modes of instability. (iv) Statistical variability will be explored through an ensemble of realizations under nominally identical conditions. (v) Simulations are currently being carried out concurrently as bundled runs using the DAKOTA toolkit. (vi) We anticipate increasingly larger petascale computational platforms to be available at the NNSA labs. (vii) We have and will continue to perform selective hero runs at super-high resolution to help quantify discretization errors to help assess the accuracy of the estimated uncertainties. (viii) UQ is being used to guide the selections of quantities to be measured with preference to those with low uncertainty, so as to avoid empty validation based on large error bars.

The Year-1 simulations of the demonstration problem employ simplified physics model: (i) a lumped detonation model, (ii) the single-phase AUSM+ flux scheme for the Euler gas equations with ideal gas equations of state, (iii) the actual particles are approximated with computational super particles, (iv) gas-particle coupling is through point-particle models of quasi-steady and unsteady forces and heat transfer, and (v) particle-particle collisions are accounted using a simplified collision model. The corresponding hero and bundled runs represent our Year-1 starting point. The above roadmap shown in Figure 1.2 lays out year-by-year progression of more detailed simulations that incorporate additional physics through new and improved models. Furthermore, each year we plan to perform larger and larger hero runs as well as large array of bundles macroscale simulations for uncertainty quantification.

The simulation roadmap is driven from the top by Uncertainty Budget (UB). A detailed phenomenon identification and ranking analysis of the demonstration problem has identified 11 key sources of errors and uncertainties which are briefly listed below:

- T1: detonation process modeling
- T2: Multiphase turbulence modeling
- T3: Real gas thermodynamic and transport properties
- T4: Inter-particle collision modeling
- T5: Particle compaction modeling (during early stages of detonation/shock propagation)
- T6: Point particle modeling of gas-particle momentum (force) exchange
- T7: Point particle modeling of gas-particle thermal (heat-transfer) exchange
- T8: Particle deformation, sintering and break-up physics
- T9: Discretization (numerical) errors
- T10: Errors from geometric approximation (geometric differences in the details of experiments and simulations)
- T11: Experimental uncertainties and measurement errors
The key activity of UB effort will be to quantify the uncertainty in the zeroth and first order prediction metrics. The zeroth order prediction metrics of the demonstration problem are:

- The blast wave location as a function of time
- The average particle front and tail locations as a function of time
- The number of large-scale instabilities of the particulate front

The first order prediction metrics go beyond the zeroth order metrics and the details of the demonstration will be quantified with the following first order metrics:

- Time evolution of the pressure at selected points within 5% error
- Time evolution of the thermal load at selected points within 20% error
- Time evolution of average particle concentration within 15% error
- Evolution of particle front thickness due to instability and turbulent mixing within 10% error
- RMS turbulent velocity and pressure fluctuations at the particle front within 15% error,
- Time evolution of local particle size distribution within 15% error
- Multiphase turbulent spectra and correlation length scales within 20% error.

An important component of the yearly UB effort is to quantify contribution from the above 11 sources of errors and uncertainties to each of the prediction metrics. This quantification will allow us to focus on error/uncertainty reduction. Thus each year we will focus on uncertainty reduction and quantification through certain modeling and simulation activities. These are the UB drivers for the proposed roadmap and they are presented at the top row of Figure 1.2.

Figure 1.2 also presents the yearly releases of CMT-nek, the new code being co-designed through an integration of exascale higher-order algorithm with exascale emulation/ simulation. Also indicated are yearly coordination with the micro, meso and macroscale experiments to be performed at Eglin Air Force Base, Arizona State University (ASU), Sandia National Laboratory (SNL) multiphase shock tube facility and Los Alamos National Laboratory (LANL) Center of Mixing Under Extreme Conditions. The macroscale simulation road map will also be supported by the yearly progression of micro and mesoscale simulations, which is also indicated in Figure 1.2.

1.4 Integration

The Center recognizes the critical importance of tight integration for the success of the center. The center will be organized in terms of tasks and cross-cutting teams, rather than in terms of faculty and their research groups. The physics-based tasks are continuous and particulates phase modeling and simulation. In addition we have exascale (EX), computer sciences (CS) and uncertainty quantification (UQ) as the cross-cutting tasks that will interface and integrate the physics-based tasks. By ensuring faculty, research scientists, and postdocs contribute to multiple physics and/or cross-cutting tasks, we will achieve tight integration. This matrix organization, depicted in Figures 1.3 and 1.4, tears down discipline and departmental boundaries and allows close interaction. In addition, significant effort has gone into integrating the various disciplines.
The intellectual integration of the different simulation and experimental talks, across the three different scales (micro, meso and macro) is shown in Figure 1.5. Uncertainty quantification, propagation and reduction along the ten sources of errors/uncertainties (T1 to T10) forms the framework that connects and drives the different simulation and experimental activities of the center. The hierarchical flow of error/uncertainty information to the macroscale is shown.
At the microscale the motion and thermal evolution of particles depends on the flow around them. In return, the particles modify the local flow by the formation of momentum and thermal wakes. Particle structures (chains and clusters) spontaneously form due to wake-wake, particle-wake and particle-particle interactions. At the mesoscale, due to inertial interaction with turbulence, particles preferentially accumulate. Also, flow instabilities can lead to large-scale structures in particle distribution. These nonuniformities have profound influence on their collective back influence on the flow. At the macroscale (or system-scale) the geometric details of the setup influence the coupling between the particles and expanding gas. Important aspects of the multiscale coupling strategy we are pursuing includes: (i) microscale-informed reduced-order descriptions (point-particle coupling models) to be employed at the mesoscale and (ii) mesoscale-informed reduced-order descriptions (multiphase LES models) to be employed at the macroscale. With this strategy, the predictive capability at the system-scale can be thoroughly validated and uncertainty rigorously quantified as illustrated in Figure 1.5.

Note that the multiscale coupling strategy and the overall uncertainty quantification plan includes both a shock-tube track and an explosive track. We have been working with the Experimental Teams at the various locations and have discussed in detail the type of characterization, inputs, and output from the experiments for a meaningful UB approach.

Finally, Figure 1.6 shows the timeline for performing the different tasks. These tasks T1-T11 were previously described.
Figure 1.6. Uncertainty Quantification Task Timeline.
2. Macroscale Team

2.1 Overview

The goal of the Macro/Mesoscale Simulation Team (MMST) is to run a problem of explosive dispersal of particles at an unprecedented level of details. Figure 2.1 provides an overview of the major phases of our demonstration problem. During year three, progress has been steady and following the trend established during year 2. Improvements on the detonation phase and the detonation products modeling of the demonstration problem have been, and continue to be, areas of substantial efforts, more analysis tools are being developed for quantifying our predictions, and simulations of the Eglin Air Force Base’s new experiments have been set up to support our model validation efforts.

2.2 Progress on the Demonstration Problem

The second generation of our demonstration problem simulation, which main results are displayed in Figures 2.2 and 2.3, was still using a lumped detonation model for the detonation phase. As initial conditions may play a significant role in the late time prediction of particles explosive dispersal, an improvement in capturing the initial energy release was an important step in improving the accuracy of our simulations. At the macroscale level, a direct calculation of the energetic material detonation is not achievable. However, a reactive burn model is desirable when accurate post-detonation conditions is sought. Using in-house-developed code RocSDT, reactive Euler simulations of PETN were run and initial profiles of key variables have been generated as new initial conditions for the demonstration problem. Next, the iterative JWL equation of state (EoS) solver was substituted for a newly developed single-equation model to warrant efficiency and stability of our simulations. This model is made of piecewise linear fits of the JWL EoS constants as functions of density and internal energy. In the limit of vanishing density and internal energy, these fits reduce the JWL equation for PETN to the ideal gas law for air.

Next, tools have been developed to characterize and quantify late time particle dispersal features, namely the number of instability waves, PM-3, and the length of instability waves PM-4. It followed that a systematic study of carefully perturbed initial particle beds has been conducted. Specifically, cylindrical detonations of particles were simulated where the initial particle volume fraction (PVF) was disturbed with a tri-modal sinusoidal perturbation as shown in Figure 2.4-left. The initial PVF then varies between 4% and 6%. Results showed on Figure 2.4-right indicate that
the particle front for the tri-modal perturbed case is faster than the case with uniform 4% PVF profile in regions with initial lower PVF (4%), and slower than the case with uniform 6% PVF in regions with high initial PVF (6%). This indicates that the fluid has a preferential path through regions with less accumulation of particles and, thus, particles in these regions are driven further by the fluid. We call this phenomenon the \textit{feedback effect}.

2.3 Code Improvement

Over the past year, progress was made in the multi-fidelity surrogate model for handling the equation of state for a detonation product and air mixture and its implementation in the center’s hydrocode, Rocflu. Two models developed last year were inserted into the code and rigorously tested. Results from these tests were compared to results from simulations using the original iterative method of handling the mixture equation of state for both tracking the blast wave and for computation time. The surrogate model showed a substantial improvement in the computation time for the overall code. As shown in Table 2.1 below, utilizing the model resulted in a speedup of a factor of approximately 2.5. On the other hand, the surrogate showed that it was not as accurate as hoped in terms of tracking certain physical phenomena. Post-processed results from the flow field showed that the model overshot the location of the primary blast wave by 15-20% at early times (the first 20 microseconds) before eventually slowing down and nearly tracking the true blast wave at later times before it escaped the domain. It did show that models constructed using the magnitude of the absolute errors of sampling points were likely to be more accurate than those constructed using both absolute and relative errors. These results showed the need to expand the variable domain of the surrogate model to include higher density values to ensure that the model does not extrapolate during the early parts of simulations. A new model was created using the new expanded domain, the same number of points as in previous models, and only the absolute error as the criteria for choosing sampling points. A surface cut of this new model is shown in Figure 2.5. Evaluation the new model is currently underway.

2.4 Collaborative Effort

In an ongoing effort towards validating point-particle models, two- and three-dimensional simulations of shock interacting with a particle curtain are being continued. Working with the UB team, different initial conditions were used to account for the experimental (Sandia MMST) uncertainty. Sensitivity studies were performed to analyze the effect of variability in volume fraction, initial thickness of particle curtain, effect of the side gap (see Fig. 2.6), and particle initial distribution on late time dispersion of particles.

Finally, simulations have been run to model the Eglin microscale and mesoscale explosive experiments with the aim of validating the current machinery governing the motion of particles in the simulations. In the microscale experiments, a single tungsten particle (d_p \sim 2 mm) was placed at the exit of a barrel that had been packed with 3 explosive pellets with detonation properties similar to LX-10-1. In the mesoscale experiments, a packet of particles with mass 0.5 grams was used in place of the single particle and one explosive pellet used instead of three. Preliminary runs modeled the explosive as a lumped system of ideal gas at high temperature with zero initial
velocity. The latest simulations use a separate equation of state to model the properties of the explosive. Figure 2.6 shows the gas density contour along with the particle position for the ideal gas case at 24 μs after complete detonation. Another experiment is expected to be performed later this year at Eglin AFB in which a packet of solid particles is accelerated in a long barrel and ejected into a compressible flow field induced by the air expelled from the barrel during particle acceleration. The geometry of the experiment invites many difficulties into the computational modeling of the experiment. Modeling strategies have been developed for this project and preliminary simulations are expected to begin before the end of the year.

2.5 Fourth Year Plans

One major objective for the MMST for year four is to perform simulations of the new set of Eglin AFB experiments: (i) the microscale gas gun, (ii) the mesoscale gas gun, and (iii) the blast pad. These experiments and their diagnostics set up have been designed with considerable inputs from the simulation team to allow for ease for future comparisons. These soon to be available high quality experimental measurements will keep strengthening and diversifying the collaboration between the MMST and the UB team.

Next, the MMST will continue working on the collision/compaction model as this model represents the last major hurdle towards a highly accurate physical modeling of CCMT’s demonstration problem. Also, the latest models from the Microscale team will be integrated in our code and tested.

Finally, work on Rocflu’s performance optimization will continue to allow for a still increasing scale resolution for the fourth generation of the demonstration problem.

<table>
<thead>
<tr>
<th>Run</th>
<th>Time steps/second (Iter.)</th>
<th>Time steps/second (Surr.)</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.389</td>
<td>18.956</td>
<td>156.54</td>
</tr>
<tr>
<td>2</td>
<td>7.273</td>
<td>19.865</td>
<td>173.13</td>
</tr>
<tr>
<td>3</td>
<td>7.325</td>
<td>19.057</td>
<td>160.16</td>
</tr>
<tr>
<td>Avg.</td>
<td>7.329</td>
<td>19.292</td>
<td>163.23</td>
</tr>
</tbody>
</table>

Table 2.1. Results showing the number of time steps per second that Rocflu computed under both the iterative method and the surrogate model (run on Vulcan using 1024 cores).
Figure 2.1. Second generation Hero-Run at late time, $t=2.5\text{ms}$. This simulation uses the iterative version of Jones-Wilkins-Lee equation of state and has an initial particle volume fraction of 40%. To prevent compaction, particles are “frozen” until the blast wave exits the initial bed of particles.

Figure 2.2. Prediction metrics for two generations of Hero-runs compared to experimental measurements Left: Blast wave location (PM-1); right: Outer front particles location (PM-2).
Figure 2.4. Left: Initial particle volume fraction (PVF) contour for $k_1 = 1$, $k_2 = 3$, $k_3 = 10$. The PVF varies from 4% until 6%. Note that $k_3 = 10$ is the one with the higher influence in the PVF distribution. The particle front for the perturbed case using $k_1 = 1$, $k_2 = 3$, $k_3 = 10$ (red line) and cases using uniform 4% PVF (blue dashed line) and uniform 6% PVF (brown dashed line) at $t=1.5\text{ms}$.

Figure 2.5. Surface slices of new, extended pressure and temperature surrogates at an explosive mass fraction of 0.5.
Figure 2.6. Effect of gap on particle dispersion. The picture has gap between curtain and the side wall. Locally the curtain spreads further than the no-gap case. A large enough concentration of particles ahead of the bulk may affect the measurement of the curtain width.

Figure 2.7. Axisymmetric simulation of Eglin AFB’s microscale gas gun experiment. Gas density contour and particle position at $t = 24\mu s$. The particle originates from the barrel exit and the simulation time begins the moment when the detonation wave has propagated through all explosives. The pure red regions show where the gas density is at least 50 kg/m$^3$. 
3. Microscale Team

3.1 Goals and Motivation

The purpose of the microscale simulations is to develop point-particle (PP) and other related models that are needed for the mesoscale and macroscale simulations of explosive dispersal of particles. These particle models are required to compute (a) instantaneous aerodynamic force on the particles, and (b) instantaneous net heat transfer between the particle and the surrounding. The commonly employed standard models for drag and heat transfer have been derived under simplifying assumptions. The objective here is to develop extended models in finite Mach, Reynolds, and volume fraction regimes – to be used in mesoscale and macroscale simulations. A strategy for a sequence of microscale simulations has been devised that will allow systematic development of the hybrid surrogate models that are applicable at conditions representative of the explosive dispersal application. The ongoing microscale simulations seek to examine particle force dependence on Mach number and volume fraction.

Over the past year, the microscale simulations team has been working on direct numerical simulations (DNS) of shocks interacting with multiple particles. The purpose of these simulations is: (a) to study the transient force and the force fluctuations experienced by the particles and (b) to apply this knowledge and use the results from simulations for developing the point-particle models.

3.2 Shock Interaction with Structured Arrays of Particles

3.2.1 Shock interaction with single particle

This is the case with zero volume fraction in which a shock travels over a single isolated spherical particle which is non-reactive and rigid. We are interested in the short time force history experienced by the particle due to the impact of the shock and for this case we can assume the particle to be stationary. Flow is assumed to be inviscid for all the cases. For these cases force history or the drag coefficient experienced by the particle is just the function of shock Mach number; see Figure 3.1. We have carried out multiple simulations, where shock Mach number ($M_s$) was varied to establish the effect of the same on the force history. For $M_s=1.22$ the post shock Mach number is less than critical (0.6), whereas for $M_s=6.0$ it is more than the critical Mach number. Hence, for $M_s=1.22$ drag on the particle goes to zero after about 20 time scales, but for $M_s=6.0$ a bow shock is formed in front of the spherical particle which results in non-zero post shock drag experienced by the particle.
Figure 3.1. Plot of non-dimensional drag (CD) against non-dimensional time (a) for Ms=1.22 (b) and Ms=6.0.

Figure 3.2. (a) Plot of non-dimensional drag coefficient ($C_D$) against non-dimensional shock time scale ($t/\tau$) Shock Mach number ($M_s$) = 6.0 and (b) Density contour plot
3.2.2 Shock interaction with one dimensional structured array of particles

Another limiting case of the shock interaction with randomly distributed particles is the shock interaction with the one-dimensional structured (horizontal) array of particles. For shock Mach number equal to 6.0, we can observe from Figure 3.2(a) that the drag increases for every successive particle in the array. There is a big jump in the drag for the second particle as compared to the first particle, and the drag levels off after the fifth particle. For this case, as seen in Figure 3.2(b), bow shocks are formed in front of each particle. There is a complex wave interaction in between the particles which results in an increase in peak drag for successive particles as compared to the previous particle. This increase in peak drag is a function of inter-particle spacing and shock Mach number. These simulations help us in understanding the dynamics of shock interaction with a structured bed of particles and to get an estimate for asymptotic increase in drag for a horizontal layer of particles.

![Figure 3.3. Three-dimensional domain for shock interaction with one layer of transverse array of particles; Ms=3.0 and 6.0.](image)

3.2.3 Shock interaction with one layer of transverse array of nine particles

We have also conducted simulations of shock interacting with one layer of transverse array of nine particles. Figure 3.3 shows the three-dimensional computational domain and results are presented in Figure 3.4. For these simulations we varied the particle spacing \( L / d_p \) as 4, 3, 2, 1.5 and 1.2 and compared the drag on the particles against the drag on a single isolated particle. We performed these simulations for \( M_s = 3.0 \) and 6.0. From these simulations we can understand the effect of neighboring particles on the drag.
Shock interaction with FCC array of particles

To study the effect of the volume fraction along with the effect of the shock Mach number, we have carried out simulations of shock interaction with particles arranged in a face centered cubic (FCC) unit cell for different shock Mach numbers and volume fractions. Table 3.1 shows the matrix of simulations and Figure 3.5 shows the three-dimensional domain. Results are presented in Figure 3.6 and 3.7. These simulations will help us understand the effect of volume fraction and shock Mach number on the drag experienced by the particles due to a shock passing over them in the presence of multiple particles.

<table>
<thead>
<tr>
<th>Table 3.1: Array of simulations for shock interacting with FCC arrangement of particles.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi = 10%$</td>
</tr>
<tr>
<td>$M_s = 1.5$</td>
</tr>
<tr>
<td>$M_s = 2.0$</td>
</tr>
<tr>
<td>$M_s = 3.0$</td>
</tr>
<tr>
<td>$M_s = 6.0$</td>
</tr>
</tbody>
</table>
Figure 3.5. a) 3-D simulation setup (b) Cut sections illustrating location of particles in the simulation domain.

Figure 3.6. Plot of non-dimensional drag ($C_D$) against non-dimensional time for $\varphi=10\%$ (a) for $M_e=1.5$, and (b) $M_e=6.0$. 
The case labelled RUN1 ($M_s=1.5$ and $\phi=10\%$) shows the same trend as the one-dimensional array, where there is increase in the successive peak drag force (see Figure 3.6a). This is because, as the shock travels over a particle it get warped and as it exits and comes out on the other side of the particle it is not planar anymore. This phenomenon is called shock focusing which leads to local pressure amplification. This local pressure amplification leads to higher peak drag force as seen in the figure. For the case RUN4 ($M_s=6.0$ and $\phi=10\%$), peak drag starts decreasing after the third plane of particles. For RUN4, the post shock flow is above critical, which results in reflected shock from the first plane of particles and formation of shocklets on the particles as seen in Figure 3.7a. Shocklets cause dissipation of energy from the lead shock causing it to weaken. This leads to reduction in peak drag force on subsequent particles. We can also observe that for all cases there is a shocklet on the last plane of particles with locally supersonic flow. This results in non-zero drag force on the last plane of particles for longer times. In the above simulations particles were not allowed to move, but if they were allowed to move the last plane of particles will move faster than rest of the bed for the cases in Figure 3.8a, 3.8c and 3.8d. This will lead to bed/particle curtain spreading observed by Wagner et. al. in their experiments.

In case of RUN1 ($\phi=10\%$ and $M_s=1.5$), the reflected is rarefaction wave, while for all the other cases it’s a reflected shock wave, indicating that the reflections for low Mach number and volume fraction are weak to form a shock wave.

In all the cases except for the first plane of particles the peak drag force on all the other plane of particles is different than that of a single isolated particle. This is the effect of volume fraction on the drag force. Currently there does not exist a micro scale model which takes into account the effect of the volume fraction and fluctuations caused by the neighboring particles on the drag force. We are currently working on this.
Figure 3.8. Contour plots of Mach number for (a) $\phi=10\%$ and $M_s=1.5$, (b) $\phi=20\%$ and $M_s=2.0$, (c) $\phi=30\%$ and $M_s=3.0$, and (d) $\phi=40\%$ and $M_s=6.0$. 
3.3 Shock Interaction with Random pack of particles

Fully resolved simulations of shock interacting with randomly distributed spherical particles were carried out with the intention of understanding the complex physical mechanisms occurring at microscale. The volume fraction was varied by changing the number of particles in the computational domain. Volume fraction was varied from 10% to 25%. Number of particles in the domain varied from 200 to 500. During early times that we are interested in, particle motion and viscous effects are negligible. Hence, we solved Euler equations on a body fitted grid with stationary particles. Simulations setup is shown in Figure 3.9. A primary shock of strength, $M_s=3.0$ was initialized. The interaction of the primary shock with the bed of particles results in a reflected shock wave traveling upstream of the particle bed and a transmitted shock traveling into the bed. The strength of the transmitted shock decreases as it travels inside the bed because of energy dissipation inside the particle bed. The rate of decrease of the transmitted shock increases as the volume fraction increases.

Shock interaction with bed of particles results in complicated wave dynamics, which can include reflected shock, transmitted shock, expansion fan, and contact interface. There is also possibility of resonance, when two wave families coincide with each other. To better understand this dynamics, we pose the problem of shock interaction with bed particles in a 1-D domain. In this 1-D domain, the particle bed is treated as a region with area change. So now it becomes a problem of shock interaction with area change. This is a Riemann problem with a discontinuous area change. We obtain the solution to this problem for different area changes (10%-25%). The wave structure from the Riemann solution is shown in Figure 3.10. The flow field resulting from the Riemann solution is plotted in Figure 3.11.

After examining the wave dynamics associated with the problem, we plot, in Figure 3.12, the inviscid non-dimensional drag coefficient, $C_D$, against the non-dimensional shock time scale, $t/\tau$ for 10%-25% volume fractions. It is clear from the drag plots in Figure 3.12 that there is significant fluctuation in the drag force seen by each particle. Also there is a clear downward trend suggesting
the decrease in strength of the transmitted shock as it travels through the particle bed. Each particle diffracts the imposed flow field that it sees, and random distribution of particles results in a unique flow field that each particle sees. This causes significant variation in the drag force. This fluctuation in the drag force is fluid-mediated and it follows a normal distribution along the varying mean drag (as a function of particle bed depth, measured from the leading edge of the bed).

Figure 3.10. Plot of approximate locations of waves in the computational domain.

Figure 3.11. Plot of normalized (a) pressure, (b) density, and (c) velocity across the domain at $t(t=12)$ for $\phi_1=10\%,15\%,20\%$, and $25\%$. 
Nonlinear effects and fluid-mediated particle-particle interaction result in transverse forces on these particles. Transverse forces are smaller compared to the stream-wise drag force but they are

Figure 3.12. Plot of streamwise drag coefficient, $C_D$, as a function of nondimensional time, $t/\tau$, for volume fractions of (a) 10%, (b) 15%, (c) 20% and, (d) 25%

Figure 3.13. Plot of peak total transverse drag coefficient, $C_{L,\text{peak}}$, for each particle. The red curve is the least squares linear fit. (a) 10%, (b) 15%, (c) 20% and, (d) 25%.
still of the same order of magnitude. Current macroscale simulations do not take into account the
effect of transverse forces as well as the particle-to-particle fluctuations in the drag force. We
compute the total transverse force, $C_L$, and its peak value for each particle is plotted in Figure 3.13
for different volume fractions. In Figure 3.13, we also plot the linear fit through the data to
understand the effect of dissipation and weakening of the transmitted shock.

Next, we plot the local Mach number and pressure contours, in Figure 3.14, to identify different
physical mechanisms that are at play during early times of shock interaction with particles. Three

Figure 3.14. Contour plot of Mach number for volume fraction 15% and $t/\tau=12$

Figure 3.15. Plot of normalized streamwise-averaged pressure for (a) 10%, (b) 15%, (c) 20% and, (d) 25%. Times shown are $t/\tau = 0$ (red), $t/\tau = 4$ (green), $t/\tau = 8$ (blue), and $t/\tau = 12$ (black).
primary mechanisms are identified – (1) Transmitted shock deforms over particles as it is traveling over particles. This results in local shock-focusing, which results in local pressure amplification and hence an increase in drag force seen by particles downstream. (2) Since the post incident shock flow is supersonic, there is formation of bow shocks and shocklets around each particle. Bow shock and shocklets dissipate energy from incident shock and weaken it. This results in a decrease in the drag force. (3) Every particle diffracts the flow field it sees, resulting in a compression or expansion wave radiating outward from the particle. Since there are multiple particles present in the particle bed, there can be either a constructive or a destructive interference between these waves and they can either increase or decrease the drag force.

Lastly, we computed the streamwise average of flow properties in the computation domain (Figure 15). The purpose of this task was to characterize the strength of the transmitted and reflected shock wave. We also made comparisons with the Riemann solution to understand the role of particles and to quantify the dissipation.

The work planned in the upcoming months includes the following:

- Perform simulations of shock interacting with bed for particles at low volume fraction and low shock Mach number.
- Perform viscous simulations to understand the role of viscosity.
- Develop and improve current point particle models to account fluid-mediated particle-particle variation.

### 3.4 Interaction of Contact Discontinuities with Particles

Shock interaction with particles has been studied extensively; however, very little is known about the effects a contact discontinuity has on particles. The primary focus of this task is the characterization of the effects of a contact discontinuity on particles. The simulations being conducted in this task consist of the following scenarios: (a) the shock and contact are initially close together, and (b) the shock and contact have an intermediate separating distance. The matrix of simulations is shown in Table 3.2, force histories on the particles are shown in Figure 3.16, and density contours are shown in Figures 3.17 and 3.18.

The current focus of this task is in the continuation of fully resolved simulations of a shock and contact interacting with a bed of particles using. The simulations being performed consist of the particles arranged in a simple cubic array, a face-centered cubic array, and a random distribution. These simulations are being run in the initial condition configurations as before (close shock and contact, and an intermediate separating distance between shock and contact) with volume fractions of 5%, 10%, 20%, and 40% and contact Mach numbers of 0.31, 0.90, and 1.26. The computational domain will be extended farther than the previous simulations so that we can capture the effects on an increased number of particles as well as the effects the contact has on later particles. Using the results of these ongoing simulations, our goal is to better understand the phenomena involved in the contact’s interaction with particles and develop models for this interaction.
Table 3.2. Matrix of simulations for the first shock followed by contact runs.

<table>
<thead>
<tr>
<th></th>
<th>$M_{ct} = 0.31$</th>
<th>$M_{ct} = 0.90$</th>
<th>$M_{ct} = 1.26$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi = 10%$</td>
<td>RUN1</td>
<td>RUN2</td>
<td>RUN3</td>
</tr>
<tr>
<td>$\varphi = 20%$</td>
<td>RUN4</td>
<td>RUN5</td>
<td>RUN6</td>
</tr>
<tr>
<td>$\varphi = 40%$</td>
<td>RUN7</td>
<td>RUN8</td>
<td>RUN9</td>
</tr>
</tbody>
</table>

Figure 3.16. Drag coefficient plotted with non-dimensional shock time scale on five particles in the three different scenarios. (a) Shock only with a volume fraction of 20% (b) Shock and contact are close with $M_{ct}=1.26$ and a volume fraction of 10% (c) Shock and contact are an intermediate distance away with $M_{ct}=1.26$ and a volume fraction of 20%.

Figure 3.17. Density contour of a central cross section of the particle bed.

Figure 3.18. 3D density contour of the particle bed.
3.5 Fourth Year Plans

Simulation matrices not completed in year 3 will be completed in year 4. Numerous archival quality papers are currently being written, and year 4 will see all year 3 work submitted for publication in archival journals. We will also continue our effort at developing point-particle force models that can be used at the macroscale.
4. Experiments

4.1 ASU Experiments

4.1.1 Goals and Motivation

Turbulent multiphase flows are highly complex. The simulation team at the University of Florida is working to simulate an exploding cylindrical charge in its entirety. This multifaced problem has been broken down into several smaller segments, including the decompression of a densely-packed particle bed.

The experimental team at Arizona State University is working to gather experimental data on the decompression of a densely-packed bed to help validate the early stage codes being developed at the University of Florida.

A particle bed composed of small (< 1mm diameter) glass beads is placed in the bottom of a vertical shock tube. A diaphragm is placed above the particle bed and everything above the diaphragm is vacuumed down to a low pressure (p₁ < 20kPa). Everything below the diaphragm remains at atmospheric pressure (p₄ = 101.3kPa). Once the diaphragm is ruptured, the pressure differential causes material (air and particles) to rapidly accelerate. The time frame of interest during each realization is on the order of milliseconds.

4.1.2 General Process

After the initial conditions, there are four primary stages of air flow and particle motion during the shocktube experiment. See Figure 4.1 for a visual description.

**Stage 0 t = 0:** All materials are essentially at rest.

**Stage 1 t = t₁:** A voltage is applied to a Nichrome wire below the diaphragm, causing the diaphragm to burst.

**Stage 2 t = t₂:** Starting at the location of the diaphragm at t = 0, a shockwave travels up and away from the particle bed while the rarefaction waves travel down towards the particle bed. Within the rarefaction waves, there is also a temperature drop.

**Stage 3 t = t₃:** Due to the rapid temperature drop, a cloud forms from the water vapor originally contained in the air above the particle bed. Additionally, the air trapped within the interstices of the top layers of the particle bed begins to escape. As this air is flowing out of the particle bed, it pushes the cloud upward and away from the particle bed and causes the particle bed to expand upward.

**Stage 4 t = t₄:** As the bed is expanding, air continues to escape the interstices and push the cloud further away from the interface. The interface between the air and the particle bed begins to deform starting with the outer edges near the glass tube.
4.1.3 Particle Image Velocimetry

Using particle image velocimetry (PIV) measurement techniques, it is possible to track movement of interfaces within the particle bed, including the rising particle/gas interface at the top of the bed and the edges of cracks/voids that develop within the bed. This technique works well in regions of the flow that have relatively low particle density compared to the initial packed bed. For this reason, PIV was used for experiments with short particle beds (initially less than 5 cm total depth). These beds become relatively dilute as the bed expands.

The PIV code used for these analyses is MatPIV (Liao & Cowen 2005). Using PIV, it is possible to better characterize instabilities as the bed rises because lateral motions can be quantified—see Figure 4.2. Additionally, there is evidence depicting a particle velocity gradient across the height of each void. As voids evolve during the expansion of the particle bed, the bottom of each void rise faster than the top, causing collapse of the voids from beneath, as pictured in Figure 4.3. Note that in both Figure 4.2 and Figure 4.3, the vertical axis is zero at the top of the tube, with the initial height of the particle bed at approximately z=28 cm.
4.1.4 Air-particle bed Interface Tracking

By tracking the evolution of the image intensity, we can characterize the deformation of the interface between the particle bed and the air above. Using $I_{\text{max}}$ as the maximum light intensity detectable ($I_{\text{max}} = 250$) for a densely-packed bed, we use three metrics to evaluate the bed location: $I=50$ (glass beads visible; denotes upper edge of the highly-disperse bed), $I/I_{\text{max}} = 60\%$ (denotes location of the average bed height), and $I=150$ (densely packed beads; denotes bottom of the disperse bed). A sample intensity
plot is shown in Figure 4.4. The full temporal record of these three metrics is shown in Figure 4.5, where spreading of the three curves is evident with increasing time. Note that each curve appears to be linear during the initial rise period. After 0.02s, the bed is relatively dilute due to the presence of large voids and interface deformation. Because the exact location of the interface becomes more difficult to define, the rise of the particle bed is difficult to measure using this technique.

![Figure 4.4: Time history of bed location, measured by relative intensity. 88-149 μm glass beads, initial bed height 32 cm, P4/P1 = 20.](image)

4.1.5 Cellular Pattern

As we are interested in quantifying the relationship between the size of the glass particles and the size of the cracks and voids generated within each particle bed, we perform a Fourier analysis to analyze the range of frequencies attainable. As particle-filled and particle-void regions show up as bright and dark regions, respectively, this analysis shows the resulting range of wavelengths, or distances between successive voids. Two sample plots are presented in Figure 4.6, which show shorter wavelengths are associated with small particles, whereas longer wavelengths are characteristic of larger particles.
4.1.6 Cloud Formation

Once the diaphragm bursts, in addition to the dramatic changes in pressure, there are also dramatic changes in temperature. Figure 4.7 shows a snapshot of the simulated temperature profile along the length of the shocktube 2.211E-4s after the diaphragm bursts provided by UF. As time goes on, the discontinuity in temperature (seen at 0.2m) will travel to the right until it is reflected at the top boundary of the shocktube. Additionally, the temperature gradient shown between 0.1m and 0.15m will expand all the way to 0 (the bottom of the shocktube), where it will be reflected upward.

Figure 4.5: Time history of bed location, measured by relative intensity. 88-149 μm glass beads, initial bed height 32 cm, $P_d/P_1 = 20$.

Figure 4.6: Fourier analysis of void size over the full temporal record of bed expansion.
Because of this dramatic drop in temperature, the water vapor in the high-pressure region of the shocktube condenses and a cloud is formed. The light intensity from this cloud is very faint. For this reason, the minimum background (the minimum value each pixel records in time) was subtracted and the color mapping was stretched significantly. See Table 4.1 for the evolution of the cloud formation.

The first frame in Table 4.1 depicts an experiment before the diaphragm has burst. An artificial image composed of the minimum value each pixel records throughout the entire experiment is subtracted from every frame. Any non-black pixels in this frame represent a low-level fluctuating background noise (flickering particles, unsteady reflections, etc.). After the first frame, very bright regions appear. These bright pixels are where large changes between the minimum light intensity and the current light intensity occur. Starting at $t = t_2$ the cloud is visible, which appears as a dimly lit cylindrical region. This is from the condensed water vapor in the high-pressure region of the shocktube. The lower the temperature, the less water vapor air can hold.
Each subsequent frame after \( t = t_2 \), shows the cloud is pushed away from the particle bed by the air escaping the interstices. It may be possible to use the cloud as a sort of gas seeding to measure the gas velocity of the air above the particle bed. Because the air escaping from the interstices pushes away the cloud, the cloud may also be used to measure the velocities of the escaped air. In addition to the cloud, the very bright cylindrical region denotes the rise and expansion of the particle bed.

4.1.7 Equipment Purchases

After getting the budget adjustment approved at the end of 2016, the team at ASU ordered the new parts for the shock tube, including a computer from Dell, a data acquisition unit from National Instruments, and pressure sensors from PCB. Together, these items will allow us to perform more highly resolved, both temporally and spatially, pressure measurements. We will also be able to trigger the high-speed camera with the data acquisition unit, meaning we will be able to synchronize our video data with our pressure measurements. This will be incredibly useful for
helping to understand the relationship between the pressure drop and the particle bed movement. Previously, this has not been possible, since the computer running the old data acquisition system was not able to run the camera software.

4.1.8 Fourth Year Plans

In 2017, the experimental team at ASU will be able to use the new equipment to better measure the pressure fluctuations that occur over time. The addition of more pressure sensors will allow for multiple measurements at the same location along the length of the shocktube. By examining multiple pressure traces from the same location along the tube, it will be possible to determine whether the one-dimensional pressure equations are appropriate to describe the flow.

A special type of data acquisition device (DAQ) was chosen to digitize the pressure measurements. The new DAQ can take over a million samples per second, which allows each of the eight new pressure sensors to be sampled at frequency of 150,000Hz. Also, the DAQ can take truly simultaneous samples. This will eliminate some distortions which are caused by only reading one channel at a time.

Additionally, with the new computer, it will be possible to synchronize the data from the pressure sensors and the high-speed camera. This is very important for better understanding gas-particle interactions.

With this new equipment, Dr. Blair Johnson will also be setting up a particle image velocimetry system. This system will be used to measure gas and particle velocities in the region above the dense particle bed.
4.2 Eglin AFB Experiments

4.2.1 Goals and Motivation

The primary goal of the experiments conducted at Eglin Air Force Base is to provide validation quality data at the micro, meso, and macroscales. AFRL/RW has completed initial experiments at the micro- and meso-scales as described in this section, and the data have been given to the UF-CCMT modeling team for uncertainty analysis. The experiments include:

a. Microscale experiments with a no-particle, detonation-only, planar geometry in support of T3, uncertainties due to thermodynamic (EOS) and transport properties;
b. Mesoscale experiments with a planar particle bed in support of T5, compaction modeling uncertainty;
c. Microscale experiments with a planar particle cluster in support of T6/T7, uncertainty in finite Re, Ma and volume fraction dependent drag and heat transfer; and
d. Macroscale experiments with an annular particle bed in support of T0, validation of the demonstration problem.

4.2.2 Microscale Experiments

Twelve small-scale explosive experiments were performed at Eglin AFB in February 2015. These experiments are considered microscale in that a small number of particles are of interest. The test data include shock arrival times, high-speed video, x-ray images, and witness panel results. The twelve experiments are equally divided into gas-valve tests (compressed helium, tests 1-6) and explosively-driven test (tests 7-12). Table salt was substituted for tungsten particles as noted in the table in an attempt to visualize the gas flow patterns with the fine particles acting as tracers.

The microscale experiments were conducted at the Advanced Warhead Experimental Facility (AWEF) at Eglin AFB, FL. All tests utilize the same basic experimental setup. The pressure array is a 3x3 grid of pressure probes, slightly offset from the test centerline with the middle probe removed due to the high likelihood of fragment impact. Four x-ray heads were used to increase the ability to track the particles’ location and velocity within the fireball.

Compressed helium provides an optically transparent shock wave and was used in the gas-valve experiments refine the settings on the diagnostic equipment, particularly the cameras. During the second day of testing, the compressed helium driver was replaced by an explosive charge. For these experiments the explosive charge consisted of three stacked N5 pellets (each 0.5” in length and diameter) initiated by an RP-83 detonator.

Data from the experimental diagnostics are detailed below. For test 8, one of the delay generators malfunctioned, resulting in a loss of Simacon camera data. The delay generator was replaced in shot 9, but an improper arming sequence resulted in a loss of all data.
The pressure probes from the compressed helium tests show a relatively weak shock (≤1 psi) and a sharp pressure rise. Alternately, the pressure probes from the explosive tests show a much stronger shock (8-23 psi) with a complex pressure rise, exhibited in a non-noise signal, followed by an oscillation, then a sharp rise in signal.

High speed images were recorded for the Phantom 6.11, Phantom Miro M310 and SIMACON. The Phantom 6.11 was used to capture the detonation products down range. A representative image from each shot for the Phantom 6.11 is shown in Figure 4.8. The SIMACON camera was used to capture the blast wave at close range, but has a faster framing rate (but limited number of images) than the Phantom Miro M310. The Phantom Miro M310 was used to capture the blast wave at close range, see Fig. 4.9.

![Figure 4.8. Representative Phantom 6.11 test images.](image)

Four x-ray heads were mounted above the particle driver. The x-ray heads were triggered on a timing delay of 10-20 microseconds; the particle velocity is of primary interest; the velocity is shown in Figure 4.11 and the x-ray images are shown in Figure 4.12. The large number and small size of the salt particles in test 10 and 11 precluded accurate velocity measurements.

Thin aluminum witness panels were used to determine the particles’ far-field impact locations in tests 7, 8, 9, and 12. The witness panels were placed 66” from the particle driver, and the center of each panel was determined using a bore sighting laser. No particles were located or recovered.
4.2.3 Mesoscale Experiments

Twenty-two mesoscale explosive experiments were performed at Eglin AFB in October-November 2015. The diagnostics and setup are the same as the microscale experiments. The 22 experiments are divided into gas-valve tests (tests 1-12) and explosively-driven tests (tests 13-22). The first tests were completed with a coarse steel powder. Tungsten powder was used for the remaining tests, where the best configuration of Tungsten required some experimentation, as seen in tests 8, 9, and 10. It was determined that the Tungsten powder was ejected most consistently when pressed in alcohol between sheets of tissue; tests 10-22 used this method.

4.2.4 Summary

The micro and mesoscale experiments performed at Eglin AFB, FL, present the UF CCMT staff with a first-look at diagnostic and measurement techniques. The ultimate objective is to provide the UF-CCMT with high quality data for validation and verification of their turbulent multiphase models.
5. UB Team

5.1 Summary of Achievements

a. Validation and UQ of the mesoscale shock tube simulation
   a. Found that inconsistency due to 3D effects between the experiment and the one-dimensional simulation is the biggest uncertainty in the validation of the one-dimensional simulation
   b. Carried out preliminary sensitivity study of individual particle force models and found quasi static force is the most influential force model contributor to model error in the mesoscale simulation
   c. Found numerical noise in particle force calculation and the finding initiated particle force model improvement
   d. Found numerical noise in volume fraction calculation and implementation of the particle based volume fraction calculation scheme resolved the issue
   e. Found that the initial volume fraction profile model is influential
   f. Carried out preliminary UQ of the two-dimensional simulation modeling the gap effect
   g. Prepared a validation framework including UQ, UB, UR and ER

b. UQ of Eglin experiments
   a. For the microscale experiments, flame from the explosion significantly hindered observation of particle and shock behaviors and it leads large uncertainty in measurements of those quantities
   b. For the microscale explosion experiment, time leg between actual detonation and sending signal was one of the most influential uncertainty sources
   c. We are planning to use a gas gun instead of explosion for the mesoscale experiment to reduce uncertainties
   d. Identified uncertainties in the initial condition of the meso/macro scale Eglin experiments

c. Validation and UQ of the BE emulation
   a. Designed experiment for validation and UQ of BE-emulation using CMT-nek, CMT-bone and CMT-bone-BE
   b. Established a framework to identify error in BE-emulation using a discrepancy between the BE-emulation and CMT-bone
   c. Carried initial investigation of noise in computation time.
   d. Applied multi-fidelity surrogate in order to combine emulations and actual computation time measurements.

5.2 Overview

The primary objective of the UB team is to estimate the model error in compressible multiphase turbulence (CMT) simulations to assess their prediction capability. Because of the presence of uncertainty in both simulations and experiments, the model error estimate involves uncertainty and uncertainty quantification (UQ) is essential for validation. However, uncertainty is also a measure
of usefulness in the model error estimate that a model error estimate with large uncertainty has little values. The unique feature of our UQ efforts is to consider uncertainty reduction (UR) in the model error estimate that would allow meaningful validation of simulations. To efficiently reduce uncertainty, uncertainty budget (UB) was calculated for reducing uncertainty of individual uncertainty sources and they been systematically reduced. We are finalizing UQ of the one-

dimensional shock tube simulation. We have found that uncertainty in inconsistency between the simulation and the experiment is the biggest uncertainty source and uncertainty in the measured initial volume fraction, which is a parameter defining initial condition of the experiment, is the second largest uncertainty source. To reduce the first uncertainty source, the center initiates a transition to two-dimensional shock tube simulation for the mesoscale.

The validations of the center are based on our UQ and validation framework prepared from experience with the one-dimensional shock tube simulation. UQ of the microscale Eglin Air Force Base experiments was carried out and meso and macroscale experiments are being planned. The validation of the microscale simulation is being carried out based on the established UQ and validation framework. With limited resource, we set the number of repetitions for experiments for UQ. To control the uncertainty due to imperfection of particle casing, we decided to fabricate the casing with designed imperfections to control the effects due to imperfection by inducing and magnifying the imperfection. Statistical properties of particles and initial volume factions were precisely identified by using CT scans and SEM images.

A secondary objective is to help other teams quantify and reduce the uncertainty in their applications. Extensive efforts are required for uncertainty identification and uncertainty quantification (UQ) effort, which poses computational challenge because UQ requires computational resource one order higher than simulation alone. To address the challenge the UB team has been evaluating and developing a surrogate based UQ framework. With the exascale team, validation of CMT-bone and CMT-bone-BE are being carried out.

Figure 5.1. Schematic figure of the shock tube (a) and particle curtain in test section (b).
5.3 Validation, Uncertainty Quantification and Uncertainty Budget of Mesoscale Shock Tube Simulation

As the CCMT simulation is composed of multiple physics, we validate the physics models through a hierarchical validation pyramid. The mesoscale shock tube simulation validation is to validate particle force model and collision model. We are wrapping up the mesoscale simulation validation and UQ, the shock tube with a particle curtain. Figure 5.1 (a) shows the configuration of the experiment where a shock is created from high pressure gas in the driver section into the driven section by bursting a diaphragm between two sections. The shock interacts with a particle curtain in the test section as is illustrated in Figure 5.1 (b). The prediction capability of the shock tube simulation is evaluated by comparing predicted/measured motion of the particle curtains; in terms of the upstream and downstream front positions of a particle curtain. Since the initial particle positions were not measured from the experiment, particle positions cannot be deterministically defined in the simulation. Average front positions of the particle curtain over the initial particle position variation are used as the prediction metrics (PM) to be compared to measurements from the experiments.

Figure 5.2 shows our validation process including UQ, UR based on UB and error reduction (ER). The framework is composed of three steps: 1) Carrying out UQ, 2) Evaluating the quantified
uncertainty whether it is large compare to the model error estimate so that UR is required. If UR is required, UR plan is implemented based on UB for efficiency and 3) Assisting the physics team to identify physical/numerical models and experiment design features causing the model error and ranking them for efficient ER.

The one-dimensional shock tube simulation is in the step of the error reduction and UQ and UR based on UB have been finished. The physics team and Sam Nili are leading efforts for reducing model error. Sam is particularly devoted to rank force models for calculating particle force in terms of their contributions to the model error using global sensitivity study. A later separated section is allocated for providing detailed information about his work.

Figure 5.3 shows the framework for UQ and model error estimation. For this framework, we assume that there is little numerical error in the simulation after verification. That is the discrepancy between predicted PM and measured PM is composed of model errors and negligible numerical error. As a first step in uncertainty quantification (UQ), we identified and classified uncertainty sources then key uncertainty sources were selected considering their effects on PM. By quantifying the uncertainty in the model error, error bar of model error estimate is obtained.
The model error of the one dimensional simulation code was estimated with the UQ and model error framework. Figure 5.4(a) shows predicted and measured PMs with their uncertainties of 95% confidence in terms of time. The width of the band represents the uncertainty in the simulation side and the band says that the true edge locations will be within the bands with 95% probability. The variations between symbols from measured PMs represent the uncertainty in experiments. Figure 5.4(b) shows the discrepancy between calculated PMs with UQ. The width of the band represents the uncertainty in the validation process. The most influential uncertainty source is the uncertainty due to inconsistency between the experiment and the shock tube simulation and the second largest uncertainty source is the uncertainty in initial volume fraction measurement in the experiment.

![Graph](image)

**Figure 5.4.** A comparison between predicted and measured particle front positions (1D). (a) Predicted and measured PMs with uncertainties of 95% confidence. (b) Prediction errors and quantified uncertainties of 95% confidence.

This results newly include uncertainty due to inconsistency. Figure 5.5(a) shows the differences between with and without considering the gap effect through two-dimensional simulation. The blue line represents measured downstream front position with the gap effect and the dashed green line represents that without the effect. Since one-dimensional simulation cannot consider the effect, it has to be compared to the measured PM unaffected. However, as Fig. 5.5(b) shows, the slit for forming a particle curtain in the test section makes impossible to observe the curtain movement from the top. Images from the side windows are available to measure the front positions and. Thus the measured PM reflects the effect of gap while the one-dimensional simulation cannot consider the effect. The uncertainty due to the inconsistency has to be included. The uncertainty can be reduced by reducing the gap effect from the shock tube experiment or incorporating it in the shock tube simulation.

Figure 5.6 shows a comparison between the two-dimensional simulation modeling the gaps and the experiments and the corresponding model error estimation with uncertainty. Except the
geometrical model, all the other physical models are identical to that used for the one-dimensional simulation. An interesting observation from Fig. 5.6(a) is that the uncertainty due to inconsistence was removed by modeling the gaps but the model error was significantly increased. This tells that the inconsistency made us underestimating the agreement last year that the disagreement between simulation and experiment was apparently small. A lesson learned from the results is that ignoring inconsistency can be the major threat that misguided us to overestimate prediction capability. The model improvement also reduced uncertainty that is reflected in the bandwidth reduction of the downstream front location.

A graduate student, Sam Nili, studies one-dimensional shock tube simulation with help of Bertrand Rollin who is in charge of macroscale simulations and Angela Diggs who was in charge of the

Figure 5.5. The effect of gap between the particle curtain and the wall. (a) Differences between w/ and w/o considering the gap effect (b) Top view of the gaps in the test section.

Figure 5.6. A comparison between predicted and measured PMs with the modified model (2D with gap). (a) Predicted and measured PMs with uncertainties of 95% confidence. (b) Prediction errors and quantified uncertainties of 95% confidence.

A graduate student, Sam Nili, studies one-dimensional shock tube simulation with help of Bertrand Rollin who is in charge of macroscale simulations and Angela Diggs who was in charge of the
one-dimensional mesoscale shock tube simulation. Justin Mathew, who is recently joined to the center, is collaborating with Justin Wagner in Sandia National Labs to identify and understand uncertainty sources in shock tube experiments for two-dimensional simulation. This interaction is now extended to include another.

5.4 Uncertainty Quantification (UQ) of Eglin Experiments

This year saw a significant increase in communication between Eglin Air Force Base (Eglin) and UF. First, Eglin was responsive to numerous inquiries regarding the experimental setup and details. This allowed identification of the major sources of uncertainty in those experiments already performed. These major sources of uncertainty are in the process of being quantified a posteriori. Of the two experiments already performed, six tests have been identified as being of interest in Table 5.1. Second, Eglin and UF collaborated in the planning of two future sets of experiment slatted to be performed in Spring 2017. This collaboration has allowed for a qualitative reduction of uncertainty before the experiments are performed, as well as identification of the major sources of uncertainty before the test. The future experiments will then have the advantage of having the uncertain inputs quantified a priori.

| Table 5.1: Six test shots of interest between the October 2014 and February 2015 Eglin microscale experiments. All tests use the same explosive |
|-----------------|----------------|
| Test #          | Particle(s)   |
| Oct14-1         | Single        |
| Oct14-2         | Ring of 7     |
| Oct14-3         | 4 x 4 Grid (16 total) |
| Feb15-1         | Single        |
| Feb15-2         | Single        |
| Feb15-3         | Diamond (4 total) |

UQ of the Eglin microscale experiments, those experiments performed in October 2014 and February 2015, have seen significant progress. As can be seen in Figure 5.7(a), particle diameter and density are identified as major sources of uncertainty. The particles are reported by the manufacturer as having nominal diameter of 2 mm and a density of 17 g/cm$^3$. The particle diameter was checked by two users independently measuring a sample of 52 tungsten spheres with Vernier calipers. The resulting distribution is shown in Figure 5.7(b) with a fitted Weibull distribution. All 52 samples were then placed into a helium gas pycnometer to measure the density. The process was repeated twelve times to give a result of 15.5±0.25 g/cm$^3$. The discrepancy in the density is under-investigation but it is hypothesized to be due to the material being a tungsten alloy. Several of the inputs, such as the explosive parameters, still need to be quantified.

Furthermore, the chosen prediction metrics of particle position and contact line have been quantified in Figure 5.8. The particle position was obtained from the x-ray images, with the timing uncertainty being quantified from oscilloscope records. The position uncertainty was obtained from the variance of the repeated tests at 30 µs and applied to all the tests. The error bars displayed
represent one standard deviation. A preliminary simulation using mean values for the uncertain inputs has been performed by Joshua Garno of the mesoscale team and shows encouraging results. The contact line has been quantified only for two tests. The other tests had insufficient frames capturing the explosive product front.

Figure 5.7. a) Expansion of the validation framework for the Eglin microscale experiments with major sources of input uncertainty identified, chosen prediction metrics, and examples of measurement uncertainty. b) UQ of particle diameter based on 52 provided sample particles from Eglin.

Figure 5.8. a) Particle position obtained for a single particle obtained from x-ray imagery. b) Contact line results obtained from high-speed Simacon.

There is significant work left to be done to bring the experiments and simulation together to perform validation. The simulation lacks many of the experimental details still, especially
regarding the initial conditions provided by the explosive. However, once the simulation has implemented a JWL equation of state and a 1D reactive burn, a surrogate will be constructed to begin propagation of the input uncertainties. The prediction metrics may then be compared and a validation metric constructed. The quantification of uncertainties is key for revealing the model error for the drag models.

Furthermore, a similar process will be applied to the mesoscale explosive tests already performed in November 2015. Quantification of the uncertain inputs has already begun. Quantification of the particle diameter has been obtained both a particle sizer and through image analysis of SEM images. The particle density has been obtained from the helium gas pycnometer. Some of the lessons learned from the microscale, such as expertise with characterization equipment, should increase the speed at which the UQ is performed.

Finally, some space will be devoted to the planning of future experiments and qualitative uncertainty reduction. Two sets of experiments are planned for Spring 2017.

1) Macroscale experiment involving a cylindrical charge with a surrounding bed of particles
2) Gas gun tests involving the acceleration of a sabot holding a bed of particles followed by the ejection of these particles

Significant discussion between Eglin and UF has been focused on performing these experiments in such a way that they fall within present simulation capabilities. For instance, in the gas gun tests the separation of the sabot and particles is not a current capability. Instead, the particles must be injected into the simulation. Therefore, the particles and sabot were chosen to be separated as close as possible to the exit of the gas gun. Another example is the effect of the explosive casing. While the literature suggests the casing may be a possible driver of instabilities in the dispersal of particles, pains were made to minimize the casing due to the inability of the simulation to capture such physics. In addition, scoring of some of the casings has been performed to attempt to control where the casing will break. Next, the new experiments will have the advantage of having the UQ of the inputs performed before the execution of the experiments. This will remove any suspicion of bias introduced from batch-to-batch variation in particle or explosive properties.

5.5 Optimization as a tool to explore the physics in particle jet formation during explosive dispersal of solid particles.

It is known that dense layers of solid particles surrounding a high energy explosive generate jet-like structures at later times after detonation. Conjectures as to the cause and subsequent development of these jet structures include: (i) imperfections in the casing containing the particles, (ii) inhomogeneities in the initial distribution of particles, (iii) stress chains within the particle bed during shock propagation and (iv) non-classical Rayleigh-Taylor and Richtmyer-Meshkov instabilities of a dispersing cloud of particles.

We have hypothesized that (i), (ii) and (iii) above produce initial variations within the bed of particles that develop into jets. We characterize this variation as particle volume fraction (PVF)
variation in space, which is defined as the volume of particles divided by the volume of particles and gas. Our future goal is to explore what initial PVF variation would lead to the observed jet formation through optimization with three mode sinusoidal perturbation parameters (amplitude, wavelength, and phase angle between modes) as design variables. Figure 5.9 shows a single mode sinusoidal perturbation of the PVF for the case with wavelength equal to 10, amplitude equal to $0.1\sqrt{2}$ and phase equal to 0. An initial hurdle was to select an objective function that would represent such jets. After substantial analysis and numerical experimentation, we divided the space into angular sectors and measure the ratio of the number of particles between the sector with most particles and the one with fewest particles (MaxR). Initial trials showed that we can start with the initial perturbation shown in Figure 5.9 with an initial a MaxR of 1.3, and grow it to a MaxR of 3.9 as it is shown in Figure 5.10(a).

Figure 5.9. Initial particle volume fraction (PVF) contour for a single mode sinusoidal perturbation with wavelength equal to 10 and amplitude equal to $0.1\sqrt{2}$.

Numerical experiments indicated substantial noise and led to our focused noise reduction efforts. The magnitude of the noise was determined by deviation from perfect rotational symmetry in the measured objective functions for corresponding rotation prescribed in the initial perturbation shown in Figure 5.10(b). The cause of the noise was determined to be a combination of randomness in the initial position of the particles and the use of Cartesian coordinates for a cylindrically symmetric problem. The proposed first measure to eliminate noise is to place the particles uniformly in each cell.
5.6 Convergence Study of 3D Shock Simulation (Microscale).

(Yiming Zhang leading this effort). Convergence of Microscale simulation is rather challenging due to three reasons: (1) Unstructured tetrahedral element (2) Multiple grid variables for refinement (3) Mixed order of convergence. These three factors will be mainly reflected as oscillatory convergence in asymptotic domain. We have reviewed and tested various literature regarding to oscillatory convergence and adopted the approach from Eça, L., & Hoekstra, M. (2014). The paper suggested four estimators for discretization error: $\delta_{RE} = \alpha h_i^p$, $\delta_1 = \alpha h_i$, $\delta_2 = \alpha h_i^2$, and $\delta_{12} = \alpha_1 h_1 + \alpha_2 h_2$. $\delta_1$, $\delta_2$ and $\delta_{12}$ could be adopted when Richardson Extrapolation $\delta_{RE}$ generated misleading order of convergence (e.g. $p<0.5$ or $p>2$). Four data points were a minimum for extrapolating oscillatory data. Asymptotic domain was assumed when data showed clear trend. Grid convergence index (GCI) was used to estimate discretization error at given runs conservatively.

The approach has been implemented in Matlab and applied by Microscale team. We performed convergence study for simulating drag coefficient using fixed particles. One of $\delta_{12}$, $\delta_1$ and $\delta_2$ was selected for best prediction accuracy while extrapolating drag coefficient on a specific particle. Refinement ratio of representative volume element size was constant 1.1 between four simulation runs.

![Figure 5.10](image-url)
runs. Estimated discretization error (GCI) at finest mesh were 0.6% ~ 2.7% for the provided particles. These estimated discretization errors were identified to be insignificant for physical analysis performed by Microscale team.

5.7 Assessing Usefulness of Behavioral Emulation (Exascale).
(Yiming Zhang leading this effort). UB team has been working closely with Exascale team on validation and uncertainty estimation (V&UE) of behavioral emulation versus CMT-bone-BE. UB team has formulated and applied a collective of data-driven approaches to assess the usefulness of emulation. We designed experiments for comparison study at 125 points scattered in 3D variable space (number of cores, number of elements, element size). Element size varies from [5,7,9,11,13], number of processors varies from [16,32,256,1024,8192], elements per processor varies from [8,32,64,128,256]. The samples are aligned on structural grids for convenient parametric analysis. The quantity of interest is total execution time at 125 samples repeated 3 times with 100 time steps.

There are two major observation by comparing \( f_{CMT\text{-bone-BE}} \) with \( f_{BE\text{ emulation}} \) using data visualization: (1) The accuracy of emulation depends on element size as seen in Fig 5.11. The discontinuous accuracy was due to interpolation schemes for calibration. The linear interpolation was replaced by lagrangian interpolation to improve calibration. (2) The behavioral emulation based on micro kernel benchmark suffered large variability between repeated runs \( \sim O(100\%) \). Therefore, exascale team sought to in-situ benchmark instead for BE emulation.

5.8 Force Sensitivity Analysis in numerical simulation of one-dimensional multiphase shock tube
Shock-particle interaction is an important phenomenon. Due to large length scale of the computational domain compared to the particle sizes for the one-dimensional shock tube
simulation, point particles and an abstract particle force model are often used to simulate the shock-particle interaction. After quantifying different sources of uncertainty of the 1D numerical simulation of Multiphase Shock Tube (MST) it was found that the largest source of uncertainty in the front particle motion is due to the modeling error. Since important parameters on the particle motions are unsteady hydrodynamic and collisional forces, quantifying and ranking the influence of their modeling error on the front particle location uncertainty are essential toward reducing the uncertainties. Front particle location was chosen as prediction metric for validation propose in this study.

The sensitivity of error in force models on the motion of dense particle curtain while particles interact with a planner shock wave is investigated. The concept of point particles where the motion of particles are traced in a Lagrangian framework and volume-fraction-based particle force models are often used to simulate the shock-particle interaction. Each force model can contain epistemic uncertainty where they cannot be determined as a fixed value, but their range can be estimated. A variance-based global sensitivity analysis allows us not only to rank the effect of error in each model on prediction metric but also to provide us with valuable information toward reducing these model uncertainty and optimizing the UQ allocated resources. In view of the uncertainty quantification, the errors in individual force components models were quantified and ranked by finding the sensitivity of the errors in terms of variation in prediction metric via global sensitivity analysis. This technique allows us an efficient uncertainty quantification and error prediction of force models. The global sensitivity analysis quantifies the importance of error in each force model on front particle movement variation as a function of time.

The force model calculates applied forces per unit mass of a particle which is denoted by \( f_p \) with the sum of forces as

\[
f_p = f_{\text{qs}} + f_{\text{pg}} + f_{\text{am}} + f_{\text{uv}} + f_{\text{ip}}
\]

where \( f_{\text{qs}}, f_{\text{pg}}, f_{\text{am}}, f_{\text{uv}} \) and \( f_{\text{ip}} \) represent the quasi-steady, pressure gradient, added mass, unsteady viscous and inter-particle force (particle collisional force) per unit mass, respectively. In order to quantify the effect of errors in force terms, the following perturbation is introduced:

\[
F = \left(1 - e_{\text{qs}}\right) f_{\text{qs}} + \left(1 - e_{\text{pg}}\right) f_{\text{pg}} + \left(1 - e_{\text{am}}\right) f_{\text{am}} + \left(1 - e_{\text{uv}}\right) f_{\text{uv}} + \left(1 - e_{\text{ip}}\right) f_{\text{ip}}
\]

Where \( e_{\text{qs}}, e_{\text{pg}}, e_{\text{am}}, e_{\text{uv}} \) and \( e_{\text{ip}} \) are uncertainty factors related to the corresponding force model. The total force \( F \) includes the model uncertainty of the force models. Each force model can contain epistemic uncertainty which cannot be determined as a fixed value, but their range can be estimated.

It is essential to investigate on each force model to accurately obtain their epistemic error bounds. However, since one of the purposes of this study is ranking the influence of their uncertainty on PM, the epistemic error band for all the force models assumed to be equal. The movement of front
particle location in one-dimensional planar shock tube model was investigated within ±10% error bounds for every force term to determine how 10% epistemic uncertainty in each individual force model changes the front particle location. To perform global sensitivity analysis, a surrogate model of front particle location was built using 42 random samples within ±10% error bounds for the combination of force terms using Latin Hypercube Sampling (LHS). The contribution of each error in forces on front particle movement is investigated by measuring its influence on the variance of front particle location. The sensitivity indices were obtained for the four force terms, which are related to the decomposition of the variance of front particle location (y) into terms either due to each input (xi) taken singularly (first order indices) as well as into terms due to the combined effects of more than one input.

\[ s_{x_i} = \frac{V(E(y|x_i))}{V(y)} \sum S_{x_i} \leq 1 \quad (3) \]

Where V and E represent the variance of expectation and S is the sensitivity index. The input samples used to calculate the sensitivity indices were computed via Monte Carlo simulation. For a linear system, the sum of sensitivity indices is one.

ASUM+up was used as a numerical scheme for the one-dimensional shock tube simulation. The particle volume fraction is 21% and 2000 point particles with 115 µ diameter used for the simulation. The particle curtain thickness is 2 mm. The average slowest 0.5% and fastest 0.5% UFP and DFP location considered as front particle locations.

The results presented for two different ways of computing volume fraction. (1) grid based volume fraction (2) particle based volume fraction. Figure 5.12 (a) and (b) show the time evolution of

![Figure 5.12](image)

(a) (b)

Figure 5.12. Time evolution of contribution of error in force model errors on UFP and DFP variation for particle based volume fraction (a) Upstream front location. (b) Downstream front location.
effect of global errors in forces and their contribution to the variance of particle laden motion for the grid based volume fraction approach.

According to Figure 5.12 (a), the contribution of error in the quasi-steady force in upstream front particle (UFP) movement is dominant at the later time. The effect of inter-particle force decrease as the shock propagates, although it is dominant at the early time. The coupling or interaction between the errors for UFP is negligible.

The results were showed in Figures 5.12 and 5.13 are for the case that the inter-particle force computed using grid based volume fraction. In this case, the volume fraction computed by summing or weighting the contribution of the Lagrangian particles within a given cell (zeroth order), between cell centers (first order). The volume fraction value for the Eulerian cell is stored at the cell center, and when local volume fraction is needed at a Lagrangian particle, the Eulerian volume fraction value is interpolated to the particle’s location. On the other hand, for the particle based volume fraction approach, particle volume fraction will first be obtained at each particle as a Lagrangian quantity, which then can be projected onto the Eulerian grid.

Figure 5.13 (a) and (b) show the time evolution of effect of global errors in forces and their contribution to the variance of particle laden motion at UFP and DFP locations for the particle based volume fraction approach. It shows that quasi-steady force is dominant for model error while the added mass force error contribution and its combination with quasi-steady force were observed only at the early time for UFP. The effect of inter-particle force for UFP in the particle based volume fraction approach was reduced and the contribution of the Higher Order Terms (H.O.T.) or the interaction between the errors were observed which it decreases as the particle laden moves.

For DFP, the contribution of pressure gradient, added mass force and higher order terms decreases and the contribution of error in quasi-steady force become dominant as front particle moves. As
Figure 5.13 (a) and (b) indicate, the effect of higher order terms decreases as front particle moves. Unlike the grid based approach the effect of the inter-particle force on PM variation is very small. However, this behavior could be due to the inappropriate model for particle collision force. In the other word, the effect of the inter-particle force could be underestimated. According to global sensitivity analysis, further improvement on quasi-steady then added mass and pressure gradient force models is required.

5.9 The Mesoscale shock tube experiments support

Starting his second semester as a master’s student at the CCMT, Justin Mathew, has continued his work from his previous time studying mesoscale shock tube experiments. He predominately works on the horizontal shock tube experiments conducted by Justin Wager at Sandia National Labs to obtain and characterize uncertainty sources. He has worked to replicate and document image processing methods to obtain key simulation parameters and prediction metrics to validate code.

In addition to Wagner’s work, Justin also studies shock tube experiments conducted by Heather Zunino at Arizona State University in an effort to formalize input and output parameters, as well as to compare with simulations being developed at the CCMT for her particular experiment. Efforts with this collaboration have been increased and are dynamic in the sense that new experimental protocols are developed simultaneously with simulations to better address uncertainties as they are recognized.

5.9.1 Sandia Horizontal Multiphase Shock Tube

Justin has completed his work replicating and documenting the methods to obtain volume fraction profiles from radiographs and particle front locations from schlieren images obtained from

![Graph of particle volume fraction profiles](image1)

![Graph of particle curtain front locations](image2)

Figure 5.14. (a) Particle volume fraction profiles obtained from radiographs and (b) particle curtain front locations obtained from schlieren images.
Wagner’s experiments. These parameters have been identified as a critical simulation input parameter and the key prediction metric by which to validate simulation, respectively. Fig. 5.14(a) shows the results of the image processing of the radiographs taken from 5 images while Fig. 5.14(b) shows the particle front locations obtained from schlieren images obtained from four separate experiments. Sources of uncertainty were identified from a variety of sources both experimental and from the image processing methods. Variation of these uncertainty sources were propagated through the image processing methods to obtain uncertainty estimates in parameters.

Front locations may also be estimated by the x-ray radiographs; however, because of magnification and blur spot effects their uncertainty is expected to be erroneously greater than that of the collimated imaging from the schlieren methods. Fig. 5.15 shows the initial x-ray profiles obtained from radiographs obtained from four x-ray experiment runs with the particle front locations as determined by schlieren images from four schlieren experiment runs. These characterizations can be used for code validation of simulations that model the particle curtain and shock interaction in the horizontal shock tube.

5.9.2 Arizona State University Vertical Shock Tube

Justin has been studying and devising experimental expectations as they are suggested by the simulation team. This allowed for the development of an experimental protocol that was established to facilitate discussion with the experimentalists to obtain reliable input parameters and prediction metrics, while maintaining control of the ability to estimate uncertainty from experimental set up. Recently, Justin has been working to develop a new experimental protocol to initiate new experiments that would allow for a complete uncertainty quantification of all input and output parameters used to validate the simulation. This new protocol considers the limitations of the simulation by, for example, accounting for inconsistencies in the diaphragm rupture which cannot be simulated.
6. CMT-nek Code Development Team

6.1 Overview

The understanding of variational methods that grew out of the Year 2 effort to recast CMT-nek into the weak form of the discontinuous Galerkin spectral element method (DGSEM) was critical to capability growth and code development in Year 3, which brought about the translation of primal forms of discontinuous Galerkin operators for higher derivatives into implementations for DGSEM and the selection of a scheme to consistently discretize diffusive terms in CMT-nek. A quasi-linear factorization expresses variable-property stress tensors as matrix products between primitive variables, transport properties, and gradients of the conserved unknowns. CMT-nek now includes diffusive terms from either the compressible Navier-Stokes equations or artificial stress tensors constructed to stabilize the Euler equations in the presence of shock waves and contact discontinuities. Testing and validation is just beginning, but CMT-nek now captures shock waves, a first for any part of nek5000 and a harbinger of its long-awaited replacement of RocFlu in wider CCMT use. The parallelization of particle tracking has continued to serve well on a large capability demonstration on LLNL’s Vulcan, and work began on mollification of point-particle forcing, enabling our first computations of volume fraction and two-way coupling.

6.2 Viscous terms


6.3 Artificial viscosity and shock capturing

to the residual in numerical evaluation of the evolution equation governing physical entropy (capped from above by the wave-speed viscosity that arises from first-order upwinding), enabled CMT-nek to capture its first full solution of unsteady flow in a shock tube (Figure 6.1). The discontinuous Galerkin spectral element methodology confines artificial viscosity to quadrature/interpolation points within a spectral element, enabling smoothed resolution of shock waves within individual elements (Figures 6.1 and 6.2), and validation against standard blast problems is ongoing (Figure 6.3).

6.4 Multiphase capabilities
The particle tracking has significantly evolved over the past few months. Originally, there existed computational and algorithmic hurdles for fully-coupled particle transport in CMT-nek. Part of the solution to these hurdles follows the work of Capecelatro and Desjardins (2013) (J. Comp. Phys. 238:1-31) in a “ghost particle” algorithm. This algorithm was independently developed this previous semester and implemented in CMT-nek. Ghost particles flexibly localize and duplicate particle data whose modeled forces affect the gas phase in more than one element, allowing mollifiers and other transfer functions to act across element faces. This will soon enable CMT-nek to accurately couple hydrodynamic particulate forces with the gas phase as well as compute inter-particle forces between neighboring particles. These new particle tracking algorithms have now shifted the particle tracking capabilities to be multiphase capabilities.

6.5 Capability demonstrations
Rarefaction flows over a regular array of 1000 spheres simulated on Mustang at LANL (Figure 6.3) produced unsteady inviscid added-mass forces that matched the undisturbed predictions of the generalized Faxen theorem of Annamalai & Balachandar (2016) (J. Fluid Mech. Submitted, under review) before acoustic reflections are felt upstream (Figure 6.4). Along with this, a one-way coupled bed of Lagrangian particles was simulated in a rarefaction flow on Vulcan at LLNL. This simulation was our first simulation on O(10^5) cores for either particles or gas solver, and it demonstrated that CMT-nek continues nek5000’s traditional high performance on legacy architectures. This simulation used 131,072 cores, 250 million grid points, and 20 billion particles.

Strong scaling results on Vulcan for a test flow of 1 billion particles 1-way coupled to the solution of the Euler equations on 1 billion grid points is shown for both the fluid solver and the particle integrator in Figure 6.6 for up to 131,072 cores.

6.6 Fourth Year Plans
Discontinuous Galerkin treatment of diffusion and viscosity was a significant challenge that changed our schedule. Artificial viscosity is not well-understood or widely used in spectral element methods for gas dynamics yet and we are contributing to the novel and fundamental research in this topic. All artificial viscosity and mass diffusivity coefficients are currently piecewise constant across elements; work is needed to smooth the pointwise variation in entropy residual and dealias the quadratures used to integrate viscous terms to refine and improve the convergence of the
subcell shock profiles. As entropy residual computations are better-understood in the spectral element context, and calibration of the artificial viscosity coefficients continues, migration of microscale simulations away from RocFlu and into CMT-nek will begin Q1Y4. The current development plan (Table 6.1) puts us on the path of performing the first simulation of the demonstration problem using CMT-nek by Q3Y4. Year 4 will also see our first consideration of real gas equations of state (starting with the JWL EOS model), entropy-bounded solution limiting (Zhang & Shu (2011) J. Comp. Phys. 230:1238-1248; Lv & Ihme (2015) J. Comp. Phys. 295:715-739) and particle-particle interactions.

Figure 6.1. Profiles of density, velocity and temperature for the shock tube problem in Section 3 of Sod (1978) J. Comp. Phys. 27:1-31 for the entropy viscosity method with viscosity capped at $c_{max}=0.25$ of the maximum wave speed in each element. Nx1=5. Exact solution follows unmarked lines/curves for each variable.

Figure 6.2. Cylindrical blast wave of a calorically perfect gas with $\gamma = 1.4$ initialized with ($\rho = 1, p = 1$) inside a circle of radius 0.4 and ($\rho = 0.125, p = 0.1$) outside. Left: Pseudocolor plot of pressure at $t=0.25$ showing the mesh with nx1=9 Gauss-Legendre-Lobatto quadrature points in each direction of each square element. Right: Pressure field (pseudocolor) and temperature (contours) at $t=1$. 
Figure 6.3. Mesh detail from a 69,200-element mesh of rarefaction flow over 20 rows of 50 spheres. The case had 50M grid points and took 50K core hours on Mustang at LANL to evolve the flow of a Mach 0.4 rarefaction emptying a quiescent driver section through a subsonic outflow.

Figure 6.4. Drag coefficient of spheres in different rows of the sphere array shown in Figure 6.4 vs time normalized by the undisturbed acoustic timescale. Symbols are computed from the model predictions of Annamalai & Balachandar “Generalized Faxen’s Theorem for Time-domain Force on a Sphere in Unsteady Viscous Compressible Flows.” Journal of Fluid Mechanics. Submitted, under review.
Table 6.1. List of capabilities that would be added to CMT-nek in the fourth fiscal year. Also shown are the projected completion dates and list of test cases to be used to verify the code.
7. CS Team

7.1 Overview

The research effort of our CS team has been to optimize CMT-nek in terms of performance, power and energy consumption. Towards that end the highlights of our accomplishments and ongoing research work are:

1. CMT-bone development and validation
2. CMT-bone on GPU
3. CMT-bone on hybrid CPU-GPU platforms

The results from these projects are summarized here:

7.2 CMT-bone development and validation

The new proxy application CMT-bone encapsulates the key data structures as well as compute and communication operations in CMT-nek as shown in Figure 7.1. CMT-bone fully retains the

![Figure 7.1. CMT-nek and CMT-bone work flow that shows the sequence of key compute and communication operations on various data structures.](image)
workflow of CMT-nek, however, it simplifies on the number of variables allocated and the number of computation and communication operations performed at each step of the work. The reduction in computations and communication operations are kept proportional to avoid undue bias towards an operation.

The four key CMT-nek operations modeled in CMT-bone are:

- Point computation for volume points. This includes preliminary computation of variables such as mass, energy, density, momentum at the volume grid points.
- Communication. This includes the exchange of surface data among processors that shares element boundaries.
- Computation. This includes computation of the fluxes and the spatial partial derivatives.
- Particles. This includes Lagrangian point particle tracking.

Validation Setup

We used VERITAS to compare the resource utilization of CMT-nek and CMT-bone and to study if CMT-bone resource utilization is representative of that of CMT-nek. All our experiment were run on Cab. Our plan is to use BG/Q also for the validation work in the current quarter.

Both applications were run on a single node and were scaled from 1 through 16 processes on one node of Cab. Workload was varied by changing the two CMT-bone modeling parameters, namely, grid size which is represented by polynomial order of the function to be fitted and the number of spectral elements. Polynomial order was varied from 5 through 15 and two different element sizes were considered 512 and 1000. The workload used for experiments is given in Figure 7.2, while the resource groups analyzed are given in Figure 7.3.

<table>
<thead>
<tr>
<th>Number of Elements</th>
<th>Workload Id</th>
<th>Polynomial Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>1000</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>15</td>
</tr>
</tbody>
</table>

Figure 7.2. Workload parameters used in our experiments.
VERITAS provides the user with two tunable parameters to control the complexity and sparsity of the predictive model it builds to jointly describe performance loss for both the applications. These parameters were set as 3 for complexity parameter and 5 for the sparsity parameter in our experiments. In general, the higher the complexity parameter, the more the number of resources that are used to explain performance loss and the higher the sparsity parameter the more the number of performance counters that are used to represent a resource.

Figure 7.3. Resource groups analyzed by VERITAS on Intel Xeon, where each group is monitored using several performance counters.

Figure 7.4. Analysis of point computation module. Most utilized resources by both applications are identified by VERITAS. Both CMT-nek and bone use L2 cache extensively. This shows that the module fits well in the instruction cache, and node memory is never a bottleneck for this application. For the smaller workloads (lower polynomial order), instruction TLB cache is the main reason for performance loss because for these sizes, most of the data fits in caches, and translation of virtual to physical addresses are exposed as bottle necks.
We analyzed the compute regions separately, and the results are presented in Figures 7.4. For the first region of point computation, the most utilized resources are MEM, L2 and L3. These resources are covered with about more than 80% coverage. Our paper has more details on how to interpret the results.

Overall, performance loss with scaling in CMT-nek is caused by cache misses and access to high-latency node memory. CMT-bone captures these performance characteristics well.

7.3 CMT-bone on GPU

After we ported CMT-bone on GPU, we evaluated the performance of the processing elements on Titan, and in what follows we first present our analysis on CPU performance and energy consumption followed by GPU performance and energy consumption.

1. CPU performance:
   A weak-scaling study was conducted on CMT-bone on a single node to observe performance, power and energy characteristics under different settings for a) number of spectral elements,
b) number of cores and c) polynomial order for modeling. The results are presented in Figure 1.5. As the number of elements per core or polynomial order increases, the time taken per time step increases as well. The increase in time with the number of cores is small and is due to resource bottlenecks exposed by the weak scaling, the main bottlenecks being the shared cache and node memory.

2. **CPU power consumption:**

To measure power on Titan, we performed our experiments on the first metered cabinet comprising of 96 nodes. CMT-bone was run on all the 96 nodes, but on different number of cores. Figure 7.6 presents the power results. CPU power consumption is generally steady and does not depend upon the program parameters or the number of cores used. Figure 7.7 shows the power consumption of a randomly picked test setup with time.

![Figure 7.6. CPU power consumption as demonstrated using various combinations of input parameters, namely numbers of cores, number of elements per core and polynomial order of CMT-bone. Power consumption is about constant and doesn’t depend upon how many cores are used.](image)

We now present a table that shows a comparison of performance of CMT-bone on GPU versus a multicore CPU. The table in Figure 7.8 shows timing results for 100 time steps for comparing the running times of GPU (Tesla, K40m) versus CPU (12-core, Intel Ivy Bridge, using all 12 cores) for different workloads represented by the polynomial order of approximation, the number of spectral elements and particles. The last column gives the ratio of CPU time to GPU time. We see from this table that the GPU is about 1.16 to 3.10 times faster than the CPU, depending on problem size. For a workload of 4096 spectral elements with polynomial order 7, and a total of 256000 particles, GPU finishes 2.22 times faster than the multi-core CPU.
3. CMT-bone on hybrid CPU-GPU platforms
   During the third quarter, we focused our research effort on developing and experimenting a
   hybrid version of CMT-bone on CPU-GPU platforms. Our platform is an Intel Ivy Bridge with
   8 GPU coprocessors, each a Tesla K40m. The results are summarized here.

7.4 CMT-bone implementation on hybrid platform
The main challenge of implementing CMT-bone on a hybrid platform was non-uniform
decomposition of spectral elements and the corresponding particles. We implemented a version of
code that allowed non-uniform data decomposition determined by load balancing strategies, the
since preliminary benchmarking showed that the GPU core could compute 20-25 times the number
of elements as compared to a CPU core in the same amount of time. Details of the GPU
implementation were presented in the previous quarterly report. The number of elements assigned
to each of the GPU cores were equal. So were the number of elements assigned to each CPU core.
Figure 7.9 shows data decomposition among 2 GPU cores and 4 CPU cores for a simple 2D
gameometry.

Our preliminary experimental results showed that when using more than 2 GPU cores, using
additional CPU cores is not useful. This is because the extra communication generated does not
compensate for the additional computational capability. Hence, most of the hybrid results (CPU
core +GPU cores) presented in this paper are limited to using 1 or 2 GPUs.
1. Performance, power and energy measurement

Power on our Intel platform was measured by a tool PowerGadget which is based on Running Average Power Limit (RAPL), while GPU power was measured using nvidia-smi. We will first present results from CPU, then GPU, followed by the results on our hybrid platform.

Figure 7.9. Pictorial example of non-uniform data decomposition of spectral elements with particles.

Figure 7.8. The ratio of processing times in CPU versus GPU for different workloads for 100 time steps. The CPU is a 12-core Intel Ivy Bridge, whereas the GPU is a Tesla K40m. We see from this table that the GPU is about 1.16 to 3.10 times faster than the CPU, depending on problem size. For a workload of 4096 spectral elements with polynomial order 7, and a total of 256000 particles, GPU finishes 2.22 times faster than the multi-core CPU.
Figure 7.10. Weak scaling on CPUs, with 2, 4, 8, and 12 cores. For tests on two CPUs, each socket uses 1, 2, 4 and 6 cores respectively. The numbers of particles and elements per core are given on the horizontal axis. Processing time increases linearly as the number of particles and elements per core is increased. Processing time also increases as more cores are used, due to additional communication time. The program runs faster when both sockets are used. Power consumption increases as the number of cores is increased and power consumption almost doubles as both CPUs are used. Energy consumption increases as the number of cores is increased, though energy consumption using one CPU is lower (higher) for smaller (larger) workloads.
2. CPU results

The CPU results are an evaluation of weak scaling on CPU cores, where the cores may be on a single CPU or distributed over two CPUs in our dual socket node. The workload consists of spectral elements and particles. We assume that the particles are uniformly distributed in the spectral elements and hence the number of particles per spectral element is the same.

Figure 7.10 shows the performance, power and energy results on CPU.

Time: The top chart in Figure 7.10 shows that time increases linearly as the number of elements per core is increased, while the number of particles has negligible impact on the processing time for the range of particles considered. As the number of cores is increased, the total amount of communication increases and hence increased overall time. Further, for the same number of cores, using cores from both CPUs gives us better performance than using them from a single CPU.

Power: The middle chart in Figure 7.10 shows that CPU power consumption depends on the number of cores as well as the number of CPUs used. Power when using two CPUs is roughly double the power when a single CPU is used (Please note that 4 cores on 2 CPUs implies, two cores were active on each CPU). Power consumption does not significantly depend on the application parameters.

Energy: The bottom chart in Figure 7.10 compares the energy consumption. For the same CPU setup, both power consumption and run time increases as the number of cores is increased. Hence, energy consumption increases as more cores are used. The relative energy consumption
for the same number of cores when a single CPU is used versus two CPUs, varies depending upon the workload. For smaller loads using one CPU is beneficial whereas for larger loads total energy consumption is lower when cores from both the CPUs are used.

Power-performance: Figure 7.11 shows the power performance data for 1 and 2 CPUs, as well as the Pareto front with optimal configurations from using 1 or 2 CPU processors with variable number of cores. The Pareto optimal configurations are: 12 and 24 cores on 2 CPUs and 1, 2, 4 and 8 core(s) on 1 CPU. This graph is useful in determining the best time and the corresponding configuration that should be used for a given power cap per processor. These results show that when using less than 12 cores, it is preferable to use only one CPU.

Energy-performance: Figure 7.12 shows the energy performance data for a variety of cores on 1 and 2 CPUs, as well as the Pareto front. The three Pareto optimal configurations are 12 and 24 cores on 2 CPUs and 8 cores on 1 CPU. To minimize energy requirements for 128K particles and 4096 Spectral Elements, 8 cores on 1 CPU results in the optimal configuration using around 131.5 seconds.

Clearly, the Pareto optimal configuration, in general, will depend on the problem size, but these results are representative of other similar problem sizes. Performance Energy tradeoffs can be determined systematically for a variety of problem sizes as described above. An appropriate
choice can then be made based on the desired tradeoffs required by the application on given

Figure 7.13. Weak scaling performance, power and energy on GPU as demonstrated using various combinations of input parameters for particles and elements per GPU. Processing time increases with the number of spectral elements and the number of particles in a spectral element, as well as with the number of GPUs. Power consumption also increases with the number of GPUs, but does not change with the number of spectral elements or particles. Energy consumption increases with increasing number of particles, number of elements and the number of GPUs.
3. GPU results

Figure 7.6 shows the performance, power and energy results on GPU.

**Time:** The top chart in Figure 7.13 shows weak scaling performance of CMT-bone on GPUs. Processing time overall increases with the number of spectral elements as well as the number of particles in a spectral element. The increase is linear with respect to the number of elements as well as the number of particles per GPU and the number of GPUs used. These results show that the communication overhead increases as the number of GPUs is increased. The increase is substantial for larger number of spectral elements when the number of GPUs is increased from 4 to 8.

**Power:** The middle chart in Figure 7.13 gives the average power consumption for the whole duration of the program. Power consumption increases as more GPUs are used and does not depend on the application parameters in any significant manner.

**Energy:** The bottom chart in Figure 7.13 shows that the energy consumption increases with increasing number of particles, number of elements and the number of GPUs. Energy increase is more profound as the number of GPUs is increased.

**Power-performance:** Figure 7.14 shows the power performance tradeoff upon using multiple GPUs, for different workloads. As expected, performance improves and power consumption increases as the number of GPUs is increased.

**Energy-performance:** Figure 7.15 shows the energy performance tradeoffs using multiple GPU for different workloads. For smaller workloads, energy consumption of GPUs varies only slightly with the number of GPUs and all the GPU configurations are energy efficient though performance drops as fewer GPUs are used. For the larger workloads using fewer GPUs is more energy efficient though there is a performance degradation. These results show that Performance Energy tradeoffs can be determined systematically for a variety of problem sizes. An appropriate choice can then be made on the desired tradeoffs required by the application on a given platform. This choice, in general, requires experimentation on a given platform and cannot be derived automatically.
Figure 7.14. Power performance trade-off using different number of GPUs and for different work load in terms of the number of elements and particles.

Figure 7.15. Energy performance trade-off using different number of GPUs and for different work load in terms of the number of elements and particles.
3.1. Hybrid results

3.1.1. Load balancing between CPU and GPU

The load balancing algorithm determines the fraction of elements (and corresponding particles) that should be assigned to the GPUs. The remaining elements (and the corresponding particles) are assigned to CPU cores. To determine this fraction, we develop a performance model, using linear regression, for variable problem sizes and number of cores/GPUs. This is done separately for the GPUs and CPUs. Once, the fraction is determined using the model, the load balancing algorithms non-uniformly decompose the spectral elements using the partitioning described earlier in this section.

Equation (1) gives the model obtained for CPU performance for polynomial order 7. Similarly, Equation (2) gives the performance of CMT-bone on GPU for polynomial order 7.

\[
T_{CPU} = 41.25 - 3.04 \times \text{cores} + 0.0187 \times \text{elements} + 5.24e^{-05} \times \text{particles} \quad (1)
\]

\[
T_{GPU} = 61.86 - 11.44 \times \text{gpus} + 0.0034 \times \text{elements} + 2.62e^{-05} \times \text{particles} \quad (2)
\]

In a hybrid environment both CPU and GPU concurrently process the spectral elements, and the performance of the combined system is best modeled by

<table>
<thead>
<tr>
<th>GPU Cores</th>
<th>CPU Cores</th>
<th>Equation based GPU load fraction</th>
<th>Optimal GPU load fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>0.43</td>
<td>0.38</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>0.60</td>
<td>0.51</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>0.67</td>
<td>0.57</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.77</td>
<td>0.92</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Figure 7.16. Percent of load processed by GPU cores for several configurations using the load balancing equation 4. The total number of elements to be processed is 8192. The number of particles is 256000, and the polynomial order is 7.
Thus, \( T_{eff} \) is minimum when \( TCPU = TGPU \). This equation can be used to determine the load distribution among CPU and GPU that results in the minimum running time of the whole system. Figure 7.16 show the percent of load (elements and particles) processed by GPU and the CPU cores. We get the estimated GPU load from our model. We also experimented with a large range of values for each configuration to determine the overall minimum. The results in the figure show that in most cases, the load fraction derived by the above equation is very close to the experimentally derived best fraction. Even when they differ the actual time requirements using the fraction derived from the above equation is not significantly worse than the optimal time derived experimentally.

### 3.1.2. Performance, power and energy consumption

**Time:** The top chart in Figure 7.17 shows the best achieved performance for various combinations of CPUs and GPUs. The load balancing between CPU cores and GPUs were determined using equation 4. The work load is 8192 elements and 256000 particles. Overall performance using two CPUs is better than the comparable configuration on one CPU. For example, with 8 GPUs and 8 CPU cores, it takes 46:816 seconds when the cores are all from one CPU, whereas it takes 29:543 seconds when 4 cores from both the CPUs are used.

**Power:** The middle chart in Figure 7.17 shows the power consumption for the different CPU-GPU combinations. Overall, power consumption using 1 CPU is lower than that using 2 CPUs.

**Energy:** The bottom chart in Figure 7.17 shows the energy consumption for the different CPU-GPU combinations. These results show the variation in the energy consumption is not significantly different for the various combinations. Additionally, the energy requirements when using 1 CPU is like 2 CPU for the same number of GPUs and cores used. This can largely be attributed to the fact that when using GPUs, the computation load assigned to the CPU cores is small enough that the difference between using 1 CPU or two CPUs is not significant for the overall time requirements.

**Power-performance:** Figure 7.18 shows the power performance tradeoff for different CPU-GPU combinations. The Pareto optimal curve can be used to derive the configuration that minimizes times for a given power requirement.

**Energy-performance:** Figure 7.19 shows the energy performance tradeoff for different CPU-GPU combinations. Most of the configurations on the Pareto-optimal curve correspond to using only the GPUs. The Pareto optimal curve can be used to derive the optimal tradeoffs between performance and energy.
Figure 7.17. Best achieved performance, power and energy consumption on hybrid GPU-CPU configurations for 8192 elements and 256000 particles, among different workload partitions between GPU and CPU cores. Performance using two CPUs is better than the comparable configuration on one CPU. On the other hand, power consumption using one CPU is better that the corresponding configuration using two CPUs. Energy consumption is comparable for the different CPU-GPU configurations.
Figure 7.18. Power performance tradeoff on hybrid GPU-CPU configurations for 8192 elements and 256000 particles. The labels denote #GPU- #cores, #CPU.

Figure 7.19. Energy performance tradeoff on hybrid GPU-CPU configurations for 8192 elements and 256000 particles as demonstrated using strong scaling. The labels denote #GPU- #cores, #CPU.
Conclusions: Thus, in summary we have implemented load balancing techniques for dividing the work for multiple GPUs and many cores potentially on multiple sockets. A non-uniform number of elements are assigned to different types of cores to achieve load balance. These techniques have been used to obtain optimal implementations for a hybrid platform. We have described experimental results of an implementation on a platform consisting of 8 Tesla K40m GPUs and 24 CPU cores.

Acknowledgement:

This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

7.5 Fourth Year Plans

Our target projects for the upcoming year are as follows:

1. CMT-bone on KNL
2. Dynamic load balancing in CMT-bone
   o Thermal load balancing
3. DVFS on CMT-bone on hybrid platform
4. Component power modeling
5. CMT-bone simulated on a system with multi-level main memory (MLM project)
8. Exascale Team

8.1 Overview

The Exascale research focuses on the development and application of Behavioral Emulation (BE), a coarse-grained simulation method, to enable rapid design-space exploration (DSE) of design strategies, optimizations, and parallelizations on extreme-scale systems up to Exascale. Such exploration supports the CCMT center by providing insight into how to optimize CMT-nek for potential candidate architectures before those architectures exist.

In this annual report, we will focus on the achievements in the following areas:

1. BE methods and tools
2. “Closing the loop” - application of BE methods and tools for CMT-nek DSE
3. FPGA acceleration of Behavioral Emulation

8.2 Behavioral Emulation Methods and Tools

The goal of Behavioral Emulation (BE) is to support design-space exploration for future Exascale applications and architectures by using simplified, component-based models which mimic much of the behavior of those applications and architectures. The achievements for this year, in terms of BE methods and tools, are as follows:

- Developed a robust coarse-grained BE simulator based on the Structural Simulation Toolkit (SST) from Sandia National Labs
- Performed large-scale experiments on DOE systems (Cab, Vulcan, Titan); simulated 100K+ MPI ranks (working toward million+)
- Investigated additional underpinnings of BE to improve BE modeling: uncertainty quantification (UQ), interpolation methods, energy modeling

8.2.1 BE-SST

A primary task this year was the development of BE-SST by leveraging the existing SST framework to enable parallel, large-scale simulations; and the freedom to build or reuse component models at different granularities. Table 8.1 describes the features provided by the SST framework and how our BE methods are built on top of the SST capabilities. BE-SST (and it derivatives) is now the primary simulation platform for the BE team. All simulation experiments for the upcoming AST review in March 2017 will be performed on BE-SST.

BE-SST v1.0 development was divided into four milestones. Milestone 1 added the capability of simulating a single BE Component using hard-coded configurations. Milestone 2 added simulation of multiple BE Components, but still using hard-coded configurations. Milestones 3 and 4 (completed in Nov) added the capability to perform full-scale parallel simulations based on user-specified configurations. Our 3-level simulation configuration model that separates operation and
application specifications from system configuration allows for easy creation of BE component libraries.

Table 8.1: BE-SST features

<table>
<thead>
<tr>
<th>SST capabilities</th>
<th>BE influences</th>
<th>BE-SST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel simulations</td>
<td>Parallel simulation, high speed-ups</td>
<td></td>
</tr>
<tr>
<td>Discrete event simulations</td>
<td>Discrete event simulation environment</td>
<td></td>
</tr>
<tr>
<td>Clock and event queues</td>
<td>Distributed component queues</td>
<td></td>
</tr>
<tr>
<td>Software definitions</td>
<td>Software definitions</td>
<td></td>
</tr>
<tr>
<td>Probabilistic simulations</td>
<td>Probabilistic simulations</td>
<td></td>
</tr>
<tr>
<td>Network models</td>
<td>Abstract network definitions</td>
<td>Abstract network models</td>
</tr>
<tr>
<td>Component models</td>
<td>Abstract hardware definitions</td>
<td>Abstract component models</td>
</tr>
</tbody>
</table>

Figure 8.1 demonstrates the scaling results of the BE-SST simulator's parallel performance. A system with 32768 cores and 163840 network links connected in a 5-D torus topology [8,8,8,8,8] was simulated on the HiPerGator cluster nodes at the University of Florida. The parallel BE-SST simulator displayed a good performance speed-up up until 128 cores (4 nodes) after which parallel overhead effects started to dominate.

Figure 8.1. Scalability and performance analysis of BE-SST running on several cores in parallel for a simulated system with 32768 cores and 163840 network links in a 5D Torus topology.
8.2.2 Scaling Studies

A major goal for this year was to perform large-scale experiments on DOE systems (Cab, Vulcan, Titan). Thus far we have simulated 100K+ MPI ranks using our BE simulator. Micro-kernel benchmarking model was used to obtain the simulator’s calibration data. The computational blocks were serially run on each of the machine for various element sizes (ranging between 5 and 25) having the elements per processor fixed at 100. We used OSU benchmarks for the communication and network models. The simulation points were validated up to 4K cores on Cab and 8K cores on Vulcan and Titan. Figure 8.2 shows the result of validation and prediction on titan. The distribution of one such validated and the predicted point is also compared. Going forward, we are working towards simulating over a million cores this coming year. The validation points will also be extended to more than 100K cores.

![Figure 8.2. Scaling studies on Titan](image)

8.2.3 Uncertainty Quantification

We collaborated with the UB team to analyze the noise and trends in the calibration data for BE. Over a design space of 125 experiments, we evaluated the accuracy of predictions for different interpolation schemes. Lagrange interpolation scheme produced best results. We are building multi-dimensional Kriging to further improve the accuracy for a generic compute block. The UB team developed multi-fidelity surrogate (MFS) models to reduce the discrepancy between CMT-bone-BE and BE simulation predictions by fitting BE emulation data on to CMT-bone, as shown in figure 8.3. We used RMSE as the metric for comparing accuracy. Using less than 20 data
samples, the MFS significantly reduces the error and provided a better fit to CMT-bone. This work will be extended to CMT-nek.

Our next phase of BE-SST development includes model enhancements in the interpolation methods and communication framework. Interpolation enhancements include the exploration of other accurate polynomial interpolation schemes such as Kriging and the development of a standardized interface that allows the AppBEO instructions and the operation configuration file to select the interpolation scheme that could provide more accurate results for a particular data layout. Communication enhancements include more fine grained network modeling and dynamic routing schemes that allows for improved accuracy and performance scaling.

![Figure 8.3. Fitting BE emulation to CMT-bone using MFS.](image)

### 8.2.4 Energy Modeling

With power consumption and thermal issues being a major constraint for Exascale computing, we have begun the process of integrating power, energy, and thermal modeling into BE framework. With these models, we will be able to explore Pareto-optimal tradeoffs between performance, power consumption, energy, and thermal issues during BE simulations. We will then be able to use this information to perform design-space exploration to optimize CMT-nek for different potential Exascale architectures.

In our effort to design an optimal benchmarking system, shown in Figure 8.4, we have decided to use Score-P and its plugins to instrument power and energy consumption API calls. Scalable and widely-supported, Score-P auto-instruments code, generates trace files and profiling data which can be visualized in Vampir. Score-P can auto-instrument processor-, node-, and wall-level power and energy consumption APIs. Score-P also has the potential to auto-instrument for thermal measurement. We started work on building and modifying plugins to allow power and energy
consumption. As the first step, we are focusing on Intel processors. So far, we have found Intel's RAPL to have the fastest update frequency and the highest accuracy.

![Figure 8.4. Scalable and extendable benchmarking system for energy benchmarking using Score-P.](image)

We plan to apply our benchmarking system to the new Intel’s new Xeon Phi Knight's Landing processors to characterize its energy consumption for scaling purposes. We have also researched and purchased an infrared camera (FLIR A35sc) to characterize heat generation of processors, specifically the Knights Landing. The Knights Landing has 64 cores and is on a very large die, thus we expect a large thermal distribution. We will study the relation between energy consumption and thermal generation in detail to potentially find thermal aware algorithms. Some images from this camera are shown in Figure 8.5. We have also purchased a Knights Landing which is currently being shipped.
8.3 Closing the Loop: DSE using BE Methods

One of the main goals of developing Behavioral Emulation (BE) methods is to help the center in developing CMT-nek for the future systems. This is achieved by performing algorithmic DSE of CMT-nek on notional Exascale architectures, where BE simulations are used to study the performance of the various alternate algorithms on these architectures.

Last quarter, we received a baseline version of CMT-nek code with particle tracking. Primarily, there are three particle tracking phases - `usr_particles_init`, `usr_particles_solver`, and `usr_particles_io`. `usr_particles_init` is called once at the start of the CMT-nek simulation to initialize the particle data structures. Upon initialization, it initiates interpolation and places particles across processor domain. `usr_particles_io` subroutine writes data to the output files at a frequency specified by the user (generally fairly low). `usr_particles_solver` is called every timestep and is the most expensive subroutine. It moves particles between processor domains based on their current location, calculates forces acting on each particle, and updates each particle's location.

<table>
<thead>
<tr>
<th>Application parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cores</strong></td>
<td>256, 1024, 4096, 16384, 32768</td>
</tr>
<tr>
<td><strong>Element Resolution (lx1)</strong></td>
<td>4, 8, 16, 20</td>
</tr>
<tr>
<td><strong>Elements/core (lelt)</strong></td>
<td>2, 8, 16, 32, 64, 128, 256</td>
</tr>
<tr>
<td><strong>Particles/core (lpart)</strong></td>
<td>0, 1024, 4096, 65536, 131072, 524288, 1048576</td>
</tr>
</tbody>
</table>
To identify good candidates for algorithmic DSE, we profiled CMT-nek to identify expensive subroutines/kernels. Based on discussions with CMT-nek team, we used the parameter space shown in Table 8.2 for profiling. The problem size is specified by the following parameters: \( l_{\text{part}} \) – defines the maximum number of particles per processor, \( l_{\text{el}} \) – defines the maximum number of elements per processor, \( l_{x1} \) – defines the element resolution.

We conducted profiling experiments on Hipergator and Vulcan. Since the parameter space is huge, we used our in-house benchmarking automation framework to easily generate job scripts for all combinations of input parameters to compile, and submit jobs. Figure 8.6 (a) shows the comparison of gas and particle subroutines with varying number of particles. Initial results show that majority of time is spent in particle solver subroutines. Interestingly, time spent in fluid solver is not independent on the number of the particles \( l_{\text{part}} \). Figures 8.6 (b) and (c) compare execution time.
of major particle solver subroutines. We observe that execution time is correlated to two parameters - lpart and lelt, and move_particles_inproc is the most expensive particle subroutine.

Since particle tracking stages which are most time consuming also stand to gain the most from DSE, we started by first looking into alternate particle tracking algorithms with CMT-nek team. They are currently working of implementing alternate algorithms for time integration and interpolation. Currently, we are in the process of developing AppBEOS of these two particle tracking stages of CMT-nek which includes the alternate algorithms. Going forward, we plan on evaluating the performance impact of the alternate algorithms on CMT-nek using BE methods and tools.

8.4 FPGA Acceleration of Behavioral Emulation

The BE approach is based on distributed discrete-event simulation (DES) between BE objects (BEOs). BEOs mimic the abstract behavior of applications or architectures using high-level surrogate models to predict each event’s impact on the targeted simulation objectives (e.g., performance, reliability, power). Although this abstract (coarse-grained) modeling approach allows for faster simulation than traditional cycle-accurate approaches (at the cost of decreased accuracy), simulation of Exascale-level problems with our current BE tool could take several minutes, hours, or even days to complete a single simulation. These lengthy simulation times place very practical limits on design-space exploration (DSE) and Uncertainty Quantification (UQ) efforts which often require thousands, millions, or even billions of independent simulations. To address this issue, we have been exploring an FPGA-accelerated approach based on pipelined DES with focus not necessarily on improved performance for a single simulation, but instead on increased simulation throughput. By focusing on throughput, we unlock the potential for huge performance gains when the problem under study calls for numerous independent simulations to be performed at the same time (e.g., DSE, Monte Carlo simulation).

In this approach, a custom compiler converts the MPI-parallel application and architecture specification used as input to our existing BE simulation tools (i.e., BEO configs) to a fully unrolled/expanded control data-flow graph (CDFG) representation (vertices represent each unique DES event and edges their input/output dependencies). The compiler utilizes several graph optimization techniques to manipulate the CDFG before ultimately mapping it to an FPGA pipeline. All vertex operations (i.e., surrogate functionality for vertex events) can be mapped to independent FPGA resources. By adding pipeline registers between events (identified by CDFG edge connectivity), a single simulation is able to flow through the circuit with a latency equivalent to the CDFG’s longest path. Moreover, successive simulations are able to start/complete once every cycle. This approach is depicted in Figure 8.7. Experiments with several application/architecture combinations that fit on a single Altera Stratix V S5GSMD8K1F40C2 FPGA have shown speedups up to eight orders of magnitude when compared to the current BE software simulator.
When the required resources for a given CDFG exceed that which is available for a single target FPGA (a likely assumption when considering Exascale simulation) we must employ some degree of resource sharing and/or circuit partitioning across multiple FPGAs. Although this approach has huge performance and scalability advantages over our original NGEEv1 approach (e.g., no idle resources when fully pipelined, the use of fixed/dedicated, unidirectional links between FPGAs rather than a shared/common bus capable of addressable any-to-any communication between thousands of threads), two obvious concerns come to mind when applying this approach to Exascale simulations. 1) although dedicated event resources scale linearly with both the number of threads and the length of simulation, overhead resources (e.g., delay registers required for event synchronization) may scale quadratically. 2) many of the graph optimization/partitioning techniques utilized by the compiler to manipulate the CDFG before ultimately mapping it to an FPGA pipeline are iterative and local. As the CDFG grows with increasing simulation scale, time to automatically generate the “best” FPGA-based circuit implementation for an arbitrary BE simulation becomes a real concern.
To address these two concerns, we focus on an approach that attempts to collapse the essentially 2D CDFG pipeline (number of threads by number of events per thread) into a 1D pipe. Here, the generated CDFG is partitioned into a spanning set of small linear subgraphs that we identify as “traces” and all remaining edges not included in the subgraphs are used to generate a temporal dependency list (the most straightforward example of this would be a CDFG partitioned such that all local operations per thread are collected as thread-level traces and all communication events between threads are used to generate a dependency list, although we need not restrict the subgraph partitioning to the thread-level). Each trace represents a sequence of events that must execute in succession to calculate the subgraph result. Each unique event type has a cost associated with it that is directly proportional to the resources required to instantiate event hardware. Event traces are combined or “aligned” into a single pipeline such that cost is minimized (multiple events of the same type aligned to the same pipeline location can share the same event hardware) and no temporal dependencies are violated. By adding pipeline registers between events, successive subgraph traces are able to flow through the pipeline and start/complete one every cycle. All traces must complete before a simulation completes. This approach is depicted in Figure 8.8.

![Figure 8.8. Collapsed Dataflow Pipeline.](image)

When considering the aforementioned straightforward case of thread-level traces, this approach has two major scaling advantages in the context of large Exascale simulation. One, pipelines scale sublinearly with the number of threads (with a factor of “threads” cost in performance) and two,
pipelines scale linearly with the length of simulation as before, but with the major advantage of only scaling as a single, unidirectional pipe that can be partitioned easily/predictably across any number of connected FPGAs with only minimal overhead.

Table 8.3 presents performance data for the two approaches while highlighting the advantages and disadvantages of each. Consider lines 1–3 & 7. The fully-expanded approach clearly has superior performance in terms of simulation throughput and latency, but consumes far more resources and is ultimately far less scalable. Although this might seem like a significant slow down, the previous speedups were eight orders of magnitude faster than software simulation, so we can afford to sacrifice this amount of performance to improve scalability. Compare lines 1-6 with lines 7-12. Resources scale linearly with the number of iteration for both approaches, but the collapsed approach allows for many more iterations due to its much lower base utilization. Consider lines 1, 7, & 13. The collapsed approach's resource utilization scales sublimely with the number of threads allowing for simulations far larger than previously possible with the full-expanded approach on a single FPGA. We are currently working on a CMT-bone.be example that spans across a configurable number of FPGAs in Novo-G#. With this example, we plan to test the scalability of these approaches for more than a single FPGA.

Table 8.3: performance of fully-expanded (left) & collapsed (right) dataflow pipelines for CMT-Bone-BE with varied MPI ranks and simulation timesteps on a single Stratix V S5GSMDD8K1F40C2

<table>
<thead>
<tr>
<th>Num. of MPI Ranks</th>
<th>Num. of Timesteps</th>
<th>Num. of Events</th>
<th>% Logic Utilization</th>
<th>Clk Rate (MHz)</th>
<th>Latency to First Output (cycles)</th>
<th>Mega Sims per Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 32</td>
<td>1</td>
<td>1,728</td>
<td>13 / 2</td>
<td>280 / 200</td>
<td>59 / 163</td>
<td>280 / 6.25</td>
</tr>
<tr>
<td>2. 32</td>
<td>2</td>
<td>3,456</td>
<td>27 / 3</td>
<td>335 / 205</td>
<td>113 / 287</td>
<td>335 / 6.41</td>
</tr>
<tr>
<td>3. 32</td>
<td>5</td>
<td>8,640</td>
<td>76 / 8</td>
<td>265 / 200</td>
<td>275 / 659</td>
<td>265 / 6.25</td>
</tr>
<tr>
<td>4. 32</td>
<td>10</td>
<td>17,280</td>
<td>- / 16</td>
<td>- / 210</td>
<td>- / 1,279</td>
<td>- / 6.56</td>
</tr>
<tr>
<td>5. 32</td>
<td>50</td>
<td>86,400</td>
<td>- / 72</td>
<td>- / 195</td>
<td>- / 6,239</td>
<td>- / 6.09</td>
</tr>
<tr>
<td>6. 32</td>
<td>100</td>
<td>172,800</td>
<td>- / -</td>
<td>- / -</td>
<td>- / -</td>
<td>- / -</td>
</tr>
<tr>
<td>7. 128</td>
<td>1</td>
<td>6,912</td>
<td>58 / 2</td>
<td>300 / 200</td>
<td>61 / 289</td>
<td>300 / 1.56</td>
</tr>
<tr>
<td>8. 128</td>
<td>2</td>
<td>13,824</td>
<td>- / 4</td>
<td>- / 200</td>
<td>- / 443</td>
<td>- / 1.56</td>
</tr>
<tr>
<td>9. 128</td>
<td>5</td>
<td>34,560</td>
<td>- / 9</td>
<td>- / 205</td>
<td>- / 905</td>
<td>- / 1.60</td>
</tr>
<tr>
<td>10. 128</td>
<td>10</td>
<td>69,120</td>
<td>- / 19</td>
<td>- / 205</td>
<td>- / 1,675</td>
<td>- / 1.60</td>
</tr>
<tr>
<td>11. 128</td>
<td>50</td>
<td>345,600</td>
<td>- / 91</td>
<td>- / 195</td>
<td>- / 7,835</td>
<td>- / 1.52</td>
</tr>
<tr>
<td>12. 128</td>
<td>100</td>
<td>691,200</td>
<td>- / -</td>
<td>- / -</td>
<td>- / -</td>
<td>- / -</td>
</tr>
<tr>
<td>13. 8K</td>
<td>1</td>
<td>442,368</td>
<td>- / 10</td>
<td>- / 225</td>
<td>- / 8,857</td>
<td>- / 0.03</td>
</tr>
</tbody>
</table>

*: configuration unable to fit on single FPGA

Overall, the proposed approach may provide simulation throughput that is many orders-of-magnitude faster than our current software approach; however, this gain in throughput comes at a cost. One notable limitation of the FPGA-pipelined approach is a sacrifice in analysis capabilities largely due to limited I/O bandwidth (e.g., the current BE simulator can log all intermediate event data for post-mortem analysis while the pipelined approach is limited to a handful of “monitored” events that can be logged without causing pipeline stalls). We also note that there are several
potential application behaviours that we may not be able to efficiently handle in the FPGA at the same level of granularity that is possible with software (e.g., behaviour that dynamically modifies application control flow based on event timing). One possible solution is to model the application at a higher level of abstraction such that the behaviours are no longer present. In summary, these limitations are less prohibitive in the context of large-scale DSE when the FPGA-accelerated BE simulator can be used to complement existing BE methods. For example, a designer can use the FPGA-accelerated BE simulator to rapidly prune a huge design space into a small set of promising candidates that can then be explored in more depth using the software BE-SST.
9. Deep Dives

9.1 Exascale Deep-dive

The University of Florida held a Deep Dive Workshop on Feb 3-4, 2015. The Agenda is presented below, and the presentation slides can be found on the CCMT webpage at:

https://www.eng.ufl.edu/ccmt/events/workshops/

Agenda:

Deep Dive
University of Florida
February 3-4, 2015

Current Attendee List:

Bob Voigt    NNSA HQ    rvoigt@krellinst.org
Matt Bement   LANL    bement@lanl.gov
David Daniel   LANL    ddd@lanl.gov
Dave Nystrom   LANL    wdn@lanl.gov
Maya Gokhale   LLNL    maya@llnl.gov
Martin Schulz   LLNL    schulzm@llnl.gov
Jim Ang     SNL    jaang@sandia.gov
Arun Rodrigues   SNL    afrodrri@sandia.gov
Jeremy Wilke   SNL    jjwilke@sandia.gov

S. Balachandar “Bala”    University of Florida    bala1s@ufl.edu
Alan George    University of Florida    george@hcs.ufl.edu
Rafi Hafika    University of Florida    haftka@ufl.edu
Herman Lam    University of Florida    hlam@ufl.edu
Sanjay Ranka    University of Florida    ranka@cise.ufl.edu
Greg Stitt    University of Florida    gstick@ece.ufl.edu
Tom Jackson    University of Florida    tlj@ufl.edu

Tania Banerjee    University of Florida    tmishra@cise.ufl.edu

University of Florida Students:
Dylan Rudolph    rudolph@hcs.ufl.edu
Nalini Kumar    nkumar@hcs.ufl.edu
Carlo Pascoe    carlo.pascoe@gmail.com
Kasim AlliKasim   kasimalli490@yahoo.com
Chris Hajas    chrishajas@ufl.edu
Mohammed Gadou   mgadou@ufl.edu
Michael Retherford
UF Deep dive agenda:

Tuesday, February 3, 2015

8:20     Van pickup at Hilton
8:30 – 9:00  Breakfast
9:00 – 9:30  Welcome and Deep-Dive Overview (3 Sessions)
   1. Behavioral emulation (BE): modeling & simulation/emulation methods
   2. CS issues (performance, energy, and thermal)
   3. Use of reconfigurable computing to accelerate behavioral emulation

* Each of the three deep-dive sessions is designed to be interactive: a combination of short presentations by UF and Tri-lab researchers, intermixed with discussion, demonstrations, etc.

9:30 – 11:30  Session 1: Behavioral Emulation: Modeling & Simulation/Emulation Methods
   • UF topics:
     o Behavioral characterization
     o Parameter estimation
   • Tri-lab topics:
     o Overview of FastForward 2 and DesignForward 2 (Jim Ang, SNL)
     o Multi-scale architectural simulation with the Structural Simulation Toolkit (Arun Rodrigues, SNL)

11:30 – 12:30  Lunch
12:30 – 2:00  Session 1 (continued): Behavioral Emulation: Beyond Device Level
   • UF topics:
     o Synchronization for speed
     o Congestion modeling
     o Behavioral characterization & modeling beyond device level
   • Tri-lab topics:
     o Using discrete event simulation for programming model exploration at extreme-scale (Jeremy Wilke, SNL)
     o ASC next-generation code projects (David Daniel, LANL)

2:00 – 5:00  Session 2: CS Issues (Performance, Energy, and Thermal)
   • UF topics:
     o Performance and autotuning for hybrid architectures
     o Energy and thermal optimization
     o Dynamic load balancing
   • Tri-lab topics:
     o Performance, energy, and thermal benchmarking (Jim Ang, SNL)
     o Why power is a performance issue: utilizing overprovisioned systems (Martin Schulz, LLNL)

* There will be an afternoon coffee break in this time slot

6:30     Dinner (University Hilton)

Wednesday February 4, 2015
8:20        Van pickup
8:30 – 9:00  Breakfast
9:00 – 11:00 Session 3: Use of Reconfigurable Computing to Accelerate Behavioral Emulation
   • UF topics:
     o Efficient mapping of behavioral emulation objects (BEOs) onto a system of FPGAs
     o Demo of current single FPGA prototype
     o Transitioning to multiple FPGAs
     o Challenges associated with maximizing emulation speed while maintaining scalability/usability
   • Tri-lab topic:
     • FPGA-based emulation of processing near memory (Maya Gokhale, LLNL)

11:00 – 12:00 Open discussion and planning for action items
12:00        Box lunch; transportation to airport as needed.

9.2  Multiphase Physics Deep-dive
The University of Florida held a Multiphase Physics Deep Dive Workshop on October 6-7, 2016. The Agenda is presented below, and the presentation slides can be found on the CCMT webpage at:

    https://www.eng.ufl.edu/ccmt/events/workshops/

Agenda:

Multiphase Physics Deep-dive Workshop
October 6-7, 2016 Attendee List

Georges Akki    gakiki@ufl.edu University of Florida
Subramanian Annamalai subbu.ase@ufl.edu University of Florida
Marco Arienti   marient@sandia.gov Sandia
S. Balachandar  bala1s@ufl.edu University of Florida
Ankur Bordoloi  ankur@lanl.gov LANL
Alexander Brown albrown@sandia.gov Sandia
Jesse Capecelatro jcaps@umich.edu University of Michigan
Seungwhan Chung  schung58@illinois.edu University of Illinois
Paul Crittenden  pcritte@ufl.edu University of Florida
William Dai     dai@lanl.gov LANL
Angela Diggs    angela.diggs.1@us.af.mil Eglin AFB
Timothy Dunn    dunn13@llnl.gov LLNL
Brad Durant     neoncrash@ufl.edu University of Florida
John Eaton      eatonj@stanford.edu Stanford University
Mahdi Esmaily   mesmaily@stanford.edu Stanford University
Center for Compressible Multiphase Turbulence

Giselle Fernandez  gisellefernandez@ufl.edu  University of Florida
Marianne Francois  mmfran@lanl.gov  LANL
Joshua Garno  jgarno@ufl.edu  University of Florida
Jason Hackl  jason.hackl@ufl.edu  University of Florida
Alan Harrison  alanh@lanl.gov  LANL
Jeremy Horwitz  horwitz1@stanford.edu  Stanford University
Kyle Hughes  kylethughes89@ufl.edu  University of Florida
Thomas Jackson  tlj@ufl.edu  University of Florida
Rahul Koneru  rahul.koneru@ufl.edu  University of Florida
Allen Kuhl  kuhl2@llnl.gov  LLNL
Ali Mani  alimani@stanford.edu  Stanford University
Yash Mehta  ymehta@ufl.edu  University of Florida
Chandler Moore  wcm0015@tigermail.auburn.edu  University of Florida
Brandon Morgan  morgan65@llnl.gov  LLNL
Balu Nadiga  balu@lanl.gov  LANL
Fady Najjar  najjar2@llnl.gov  LLNL
Brandon Osborne  bosborne3@ufl.edu  University of Florida
Fred Ouellet  f.ouellet@ufl.edu  University of Florida
John Parra-Alvarez  jcparraa@gmail.com  University of Utah
Katherine Prestridge  kpp@lanl.gov  LANL
Bertrand Rollin  bertrandrollin@gmail.com  Embry-Riddle
Kevin Ruggirello  kruggir@sandia.gov  Sandia
Kambiz Salari  salari1@llnl.gov  LLNL
Shane Schumacher  seschum@stanford.edu  University of Florida
Philip Smith  philip.smith@utah.edu  University of Utah
Sean Smith  sean.t.smith@utah.edu  University of Utah
Prashanth Sridharan  shan1130@ufl.edu  University of Florida
Jeff St. Clair  jeff.stclair@us.army.mil  University of Utah
Cameron Stewart  csstewart10@ufl.edu  University of Florida
Jeremy Thoemmel  jthoemmel@gmail.com  University of Utah
Gretar Tryggvason  gtryggva@nd.edu  University of Notre Dame
Markus Uhlmann  markus.uhlmann@kit.edu  Karlsruhe Institute of Tech.
Laura Villafane  lvillafa@stanford.edu  Stanford University
Robert Voigt  rvoigt@krellinst.org  ASC AST
Seng Keat Yeoh  syeoh@illinois.edu  XPACC University of Illinois
Duan Zhang  dzhang@lanl.gov  LANL
Ju Zhang  jzhang@fit.edu  Florida Institute of Tech.
Agenda Multiphase Physics Deep-Dive, October 6-7, 2016
St. Petersburg Marriott Clearwater,
12600 Roosevelt Blvd, North St. Petersburg, FL 33716, Phone: 727-572-7800

Thursday Oct 6th, 2016
7:30 – 8:30 Breakfast (Provided)
8:30 – 8:45 Welcome (S. Balachandar)
8:45 – 10:15 Overviews
   Gretar Tryggvason Challenges and opportunities in fully resolved simulations of multi fluid flows
   Ali Mani Overview of computational modeling at Stanford PSAAP: particle-laden flows subject to radiative heating
   Kambiz Salari Research activities for energetic dispersal of particles
10:15 – 10:30 Coffee
10:30 – 12:00 Overviews
   Alex Brown Multiphase Methods for Modeling Fire Environments
   Phil Smith Multi-phase flow modeling at Utah PSAAP – predictivity in application
   S. Balachandar Overview of multiphase flow computational strategy at UF PSAAP
12:00 – 12:15 Further Discussion
12:15 – 1:30 Lunch (Provided)
1:30 – 3:00 Macroscale
   Duan Zhang Equations and Closures for Deformation and Flow of Continuous and Disperse Materials
   Allen Khul 3-Phase Model of Explosion Fields
   Marco Arienti Multiphase Flow Simulation Strategies at the CRF
3:00 – 3:15 Coffee
3:15 – 5:15 Microscale
   Jeremy Horwitz Point-particle modeling for two-way-coupled problems: Challenges, verification, and physics-based improvements
   Georges Akiki Extended point particle model
   Jesse Capecelatro Recent insights on turbulence modeling of strongly-coupled particle-laden flows
   Tom Jackson Microscale simulations of shock particle interaction
6:30 – 9:00 Dinner (Provided – all attendees)

Friday Oct 7th, 2016
7:30 – 8:30 Breakfast (Provided)

8:30 – 10:30 Modeling & numerical methods

- **Sean Smith** Particle dynamics: coal-specific modeling
- **A. Subramaniam** Microscale modeling based on Generalized Faxen theorem
- **Alan Harrison** Modeling of Ejecta Particles in the FLAG Continuum Mechanics Code
- **Markus Uhlmann** Large scale microscale simulations and modeling opportunities

10:30 – 10:45 Coffee

10:45 – 12:15 Experiments and simulations

- **Ankur Bordoloi** Experimental measurements of drag on shocked particles
- **Laura Villafañe** Including real experimental effects in validation of numerical models for confined particle-laden flows
- **Fady Najjar** Meso-scale Simulations of Shock-Particle Interactions

12:15 – 1:30 Lunch (Provided)

1:30 – 3:00 Meso/macroscale

- **Mahdi Esmaily** A systematic study of turbophoresis by four-way-coupled simulation of Stokesian particles in channel flow
- **Balu Nadiga** Bayesian Analysis of Inter-Phase Momentum Transfer in the Dispersed Eulerian Formulation of Multiphase Flow
- **John Parra-Álvarez** Eulerian Models and Polydispersity Treatment for Dilute Gas-Particle Flows

3:00-4:00 Discussion and Closing remarks
10. Publications

2014


2015


2016
2017

11. Conferences and Presentations

2014

2015


2016


50. F. Ouellet, C. Park and B. Rollin, S. Balachandar, “A Multi-Fidelity Surrogate Model for
111

60. B.A. Johnson & E.A. Cowen (2016) Turbulent Boundary Layers and Sediment Suspension Absent Mean Flow-Induced Shear: An Experimental Study; American Geophysical Union, San Francisco, CA, December 2016.
63. “Optimizing CMT-nek for performance and power consumption”, by Tania Banerjee on 6/30/2016 at Sandia National Laboratories, Livermore, CA.
64. “Multiobjective Optimization of CMT-bone on Hybrid Processors” by Tania Banerjee on 11/8/2016 at Hangzhou, China.


2017


12. Workshops Held or Attended

2014
2. Dr. Tania Banerjee. Salishan Conference, April 2014.
5. Heather Zunino presented tomographic PIV at a PIV workshop to a group of faculty, post-docs, and Ph.D. students at Instituto Superior Técnico (IST) in Lisbon, Portugal. She was invited to stay for this workshop and present after meeting with several researchers from IST at the Laser Symposium this summer.

2015
10. Heather Zunino presented tomographic PIV at a PIV workshop to a group of faculty, post-docs, and Ph.D. students at Instituto Superior Técnico (IST) in Lisbon, Portugal. She was invited to stay for this workshop and present after meeting with several researchers from IST at the Laser Symposium summer 2015.

2016
12. Dr. Sanjay Kumar and Ph.D. students Kasim Alli and Carlo Pascoe (2016). Attended the WEST workshop. Dr. Kumar gave a talk.
15. Rahul Koneru and Fred Ouellet attended “Scaling your science on Mira” workshop, May 24-25, Argonne National Laboratory.
13. Students and Staff Internships

13.1 Internships Completed

2014

1. **Heather Zunino**, Ph.D. Student (ASU), US, Dr. R. Adrian. Ms. Zunino completed her 10-consecutive week stay at Los Alamos National Laboratory, under the mentorship of Dr. Kathy Prestridge this summer (May-August 2014). Dr. Prestridge is the Extreme Fluids Team leader in the Physics-23 group at LANSCE. Project: Vertical Shock Tube (Calibration and Tomographic PIV), Horizontal Shock Tube (Particle Tracking Program).

2. **Kevin Cheng**, MS Student, Florida. Lawrence Livermore National Laboratory. Mentor: Dr. Maya Gokhale, Dr. Scott Lloyd. Project: An Emulation Framework for Tracing near Memory Computation. US, Dr. Alan George, ECE, MS (graduated Fall 2014), core.


2015


5. **Dr. Bertrand Rollin**, Staff Scientist, CCMT. March 16-20, 2015, LANL.


7. **Chris Hajas**, M.S. Student, US, ECE, Dr. Herman Lam, core. May 18-August 18, 2015 at LLNL with Dr. Maya Gokhale.

8. **Christopher Neal**, Ph.D. Student, US, MAE, Dr. S. Balachandar, core. June 14-August 20, 2015 at LLNL with Dr. Kambiz Salari. [Chris Neal (graduate student) spent 10 weeks at Lawrence Livermore National Lab (LLNL) over summer 2015. He was paired with Dr. Kambiz Salari as his mentor who had similar research interests. The work that was performed will not be described here. He was asked by Dr. Salari to return to LLNL next summer for a second internship at the lab. Chris’ work will also be highlighted in the Stewardship Science Academic Programs (SSAP) Annual, which highlights the exceptional research activities and people supported by the SSAP (PSAAP is under the SSAP).]

9. **Carlo Pascoe**, Ph.D. Student, US, ECE, Dr. Herman Lam, core. Will intern summer 2015 at LLNL with Dr. Maya Gokhale.


11. **Dr. Tania Banerjee**, PhD., one week internship in May, 2015 to LLNL.

2016
13. Dr. Chanyoung Park, Postdoc, CCMT. Feb 22-26, 2016 at LLNL with Dr. Samuel P. Schofield.
14. Dr. Jason Hackl (Postdoc, CCMT) visited Sandia National Labs, Albuquerque, New Mexico from January 16-20 to present CMT-nek to Greg Weirs and other personnel.
17. Georges Akiki, Ph.D. Student, US, MAE, Dr. Balachandar, core. Internship to LANL. May-August, 2016, Dr. Marianne Francois.
18. Dr. Tania Banerjee, PhD., one week internship in June, 2016 to Sandia.

13.2 Internships Planned
1. Paul Crittenden, Ph.D. Student, US, MAE, Dr. Balachandar, core. Internship planned to LANL for summer 2017 with Dr. Marianne Francois.
2. Mohamed Gadou, Ph.D., Dr. Sanjay Ranka, Computer Science, UF, core

13.3 Internships Not Yet Planned
1. Fred Ouellet, Ph.D. Student, US, MAE, Dr. S. Balachandar, core
2. Prashanth Sridharan, Ph.D., US, MAE, Dr. Balachandar, core
3. Yash Metha, Ph.D., MAE, Dr. Balachandar, cost share (internship not required)
4. Yiming Zhang, Ph.D., China, MAE, Drs. Haftka and Kim, cost share (internship not required)
5. Rahul Koneru, Ph.D., India, MAE, Dr. Balachandar, cost share (internship not required)
6. Joshua Garno, Ph.D. Student, US, MAE, Dr. Balachandar, core
7. Brandon Osborne, Ph.D. Student, US, MAE, Dr. Balachandar, core
8. Kyle T. Hughes, Ph.D. Student, US, MAE, Drs. Haftka and Kim, core
9. Sam Nili, Ph.D. Student, US, MAE, Drs. Haftka and Kim, core
10. Brad Durant, Ph.D., MAE, Dr. Balachandar, core
11. Ryan Blanchard, BS, US, ECE, Dr. Herman Lam, core
12. Keke Zhai, Ph. D. student, China, Dr. Ranka, CISE, UF, cost share (internship not required)
13. Sankeerth Mogili, MS student, Dr. Ranka, CISE, UF
14. Ajay Ramaswamy, ECE, Dr. Herman Lam, cost share, (internship not required)
15. Aravind Neelakantan, ECE, Dr. Herman Lam, cost share, (internship not required)
16. Mason Rawson, ECE, Dr. Herman Lam, core, no visit planned yet
17. Sai Prabhakar Rao Chenna, ECE, Dr. Herman Lam, cost share, (internship not required)
18. Trokon Johnson, Ph.D. student, ECE, Dr. Herman Lam core, visit planned to LANL in summer 2017.
13.4 Graduated Students

1. **Kevin Cheng**, MS Student (2014). Dr. Alan George, ECE.
8. **Parth Shah**, M.S. Student (2016), India, ECE, Dr. Herman Lam.

13.5 Placement of Staff

1. **Dr. Bertrand Rollin**, Staff Scientist, CCMT. Faculty position at Embry-Riddle, Fall 2015.
2. **Dr. Mrugesh Shringarpure**, Postdoc, CCMT. Researcher, Exxon Mobil, Spring 2016.
14. NNSA Laboratory Interactions

2014
1. Rob Cunningham, LANL. Setting up "Florida" file sharing group on Mustang
2. Blaise Barney, LLNL. Setting up account on Vulcan and Surface.
4. Nathan Fabian, SNL. Introduction to Catalyst.
5. Don Frederick, LLNL. Issue with submitting a run on Vulcan.
7. Jan Nunes, LLNL. Account request on Edge
8. Discussions with Donald Frederick of Lawrence Livermore National Laboratory related to MPI issues on the Vulcan computing cluster – discussion was relevant to parallel operation of the Rocflu code.
9. Telecon with Paraview Catalyst developers and users (Greg Weirs, Nathan Fabian, Kenneth Moreland at Sandia National Laboratory) at about deploying the Catalyst library into Rocflu for in-situ visualization.
10. Worked with Greg Lee (Livermore Computing Center)–who is a software debugging/troubleshooting expert at LLNL–to get Rocflu to run at scale on the Vulcan computing cluster at LLNL.
11. Interactions with Rich Cook (Livermore Computing) who is the visualization expert at LLNL–we are working with him to get in-situ visualization capabilities using Catalyst integrated into Rocflu.
12. Interacted with David DeBonis at Sandia National Laboratories to get setup with using PowerInsight for power and energy measurements.
13. Interacted with Robert Cunningham and Amanda Bonnie at Los Alamos National Laboratory for temperature measurements using the data collection tool Lightweight Distributed Metric Service.
14. Interacted with Justin A. Too and Daniel J. Quinlan at the Lawrence Livermore National Laboratory on compilation and installation of ROSE.
15. Interacted with Blaise Barney at the Lawrence Livermore National Laboratory on various system issues, including access to CVS.
16. Road trip to SNL (Albuquerque) and LANL, March 24 – 26, 2014 (Herman, Bala, and Rafi)
17. SNL (Albuquerque): Jim Ang, Arun Rodrigues, Scott Hemmert, Simon Hammond - from SST team at SNL, Albuquerque, NM
18. SNL (Livermore): Jeremiah Wilke from SST Macro team at SNL, Livermore, CA

2015
19. LLNL: Maya Gokhale regarding proposal to analyze memory access demands of CMT-Nek kernels and evaluate the potential benefits of utilizing stacked memories with embedded logic functions to increase performance
20. Dr. Steve Beresh (SNL), visit CCMT and gave talk, Thursday April 23, 2015
21. The members of the microscale group (in particular Chris Neal and Yash Mehta, another graduate student) have worked closely with the staff at LLNL's Livermore Computing center to resolve an issue that Rocflu was having when it was scaled up and executed on
the BG/Q machine Vulcan. We have also been in touch with Rich Cook at LLNL's Livermore Computing to continue our work with Paraview's Catalyst library integration into Rocflu. During the fall and spring, Chris Neal has been in touch with Dr. Kambiz Salari to keep him up-to-date on the progress of Chris's research. Chris and Yash were aided by Blaise Barney (LLNL) in renewing their computing accounts at LLNL.

22. Drs. Jason Hackl and Mrugesh Shringarpure and David Zwick and Goran Marjanovic. Visited Argonne National Lab for the 2015 Nek User/Developer meeting to commit CMT-nek to the nek5000 repository and work intensively with Prof. Paul Fischer.

23. Barry Rountree, LLNL. Tania worked closely with Barry’s group to come up to speed on using RAPL to measure and bound power on Intel platform

24. Tanzima Islam, LLNL. Tania is working with Tanzima on validating CMT-bone proxy app

25. David DeBonis, SNL. Tania interacted with David to come up to speed on using PowerInsight for power and energy measurements.

26. Patrick Lavin, a summer intern with Barry Rountree, worked on the performance issue and used vectorization effectively on the derivative optimization code getting about 40% improvement using vectorization alone. Tania will follow up with verifying if it is possible to get further improvement with CHiLL based autotuning approach.

27. S. Balachandar and H. Lam attended and presented at the 2015 ASC PI meeting at Monterey, CA.


2016

29. S. Balachandar and S. Ranka attended and presented at the 2016 ASC PI meeting at Las Vegas, NV.

30. Tania Banerjee collaborated with Jeremy Wilke, Gwen Voskuilen and Arun Rodriguez on MLM.

31. Tania Banerjee collaborated with Tanzima Islam to validate proxy app CMT-bone.

32. Justin Matthews and Chanyoung Park collaborated with Justin Wagner (Sandia).

33. Kyle Hughes collaborated with Kathy Prestridge (LANL) in sharing p-rad experiments of rapidly dispersed pre-frAGMENTED particles.

34. Jeff St. Clair and Balachandar are interacting with Fady Najjar (LLNL) in performing ALE3D simulations of intense shock propagation over a close-packed bed of deformable aluminum particles.

35. Balachandar interacted with John Schwarzkef (LANL) in completing a review article on compressible multiphase flow for the second evolution of multiphase flow handbook.

36. Balachandar and Jackson organized and hosted the Deepdive workshop on multiphase flows. Approximately 15 research scientists from the three NNSA laboratories attended.

37. Yash Metha interacted with Kambiz Salari (LLNL) on three-dimensional simulations of a strong shock propagating through a random bed of particles. This work has been written up and is in the final stages of submission as an archival journal paper.
38. Tom Jackson, Fady Najjar (LLNL), and Habib Najm (Sandia) are organizing a Focus Topic session on UQ in high speed flows for the upcoming APS SCCM meeting in July.
39. The UQ team has regular teleconference calls with Greg Weirs (Sandia).