CCMT

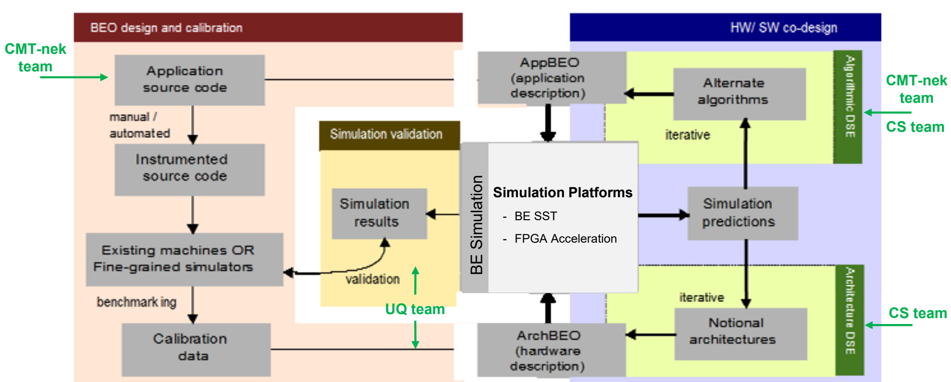
BE-Simulation Framework: Looking Under the Hood

Nalini Kumar, Ph.D. Candidate

Center for Compressible Multiphase Turbulence (CCMT)
ECE Department, University of Florida

UF UNIVERSITY of FLORIDA

Co-Design Using *Behavioral Emulation*



* BEO – Behavioral Emulation Object

HW/SW co-design

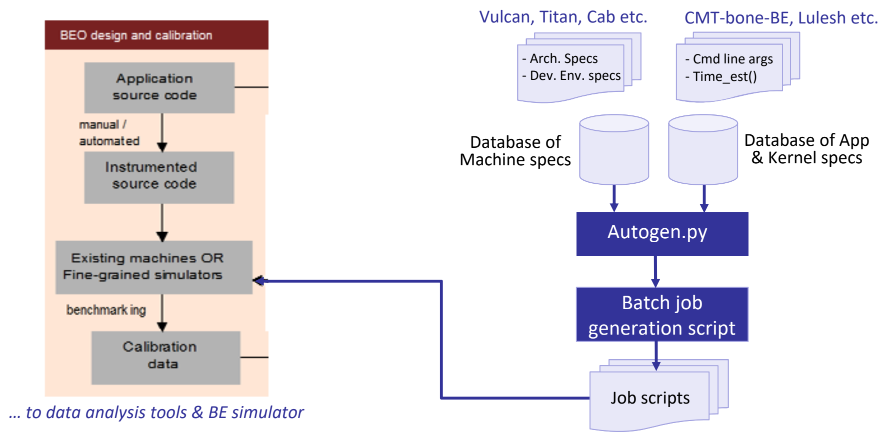
- Algorithmic & architectural design-space exploration

Coarse-grained BE simulation

- Balance of simulation speed & accuracy for rapid design-space evaluation

CCMT | 2

BE Framework Automation



- Designed a framework to simplify and automate benchmarking
  - Job scripts can be generated for different machine/app combinations
  - Easy to extend benchmarking to new applications & systems
- Leveraging similar framework for CMT-nek batch runs as well

CCMT | 3

Application to AppBEO Conversion

Original Application → **Call Graph**

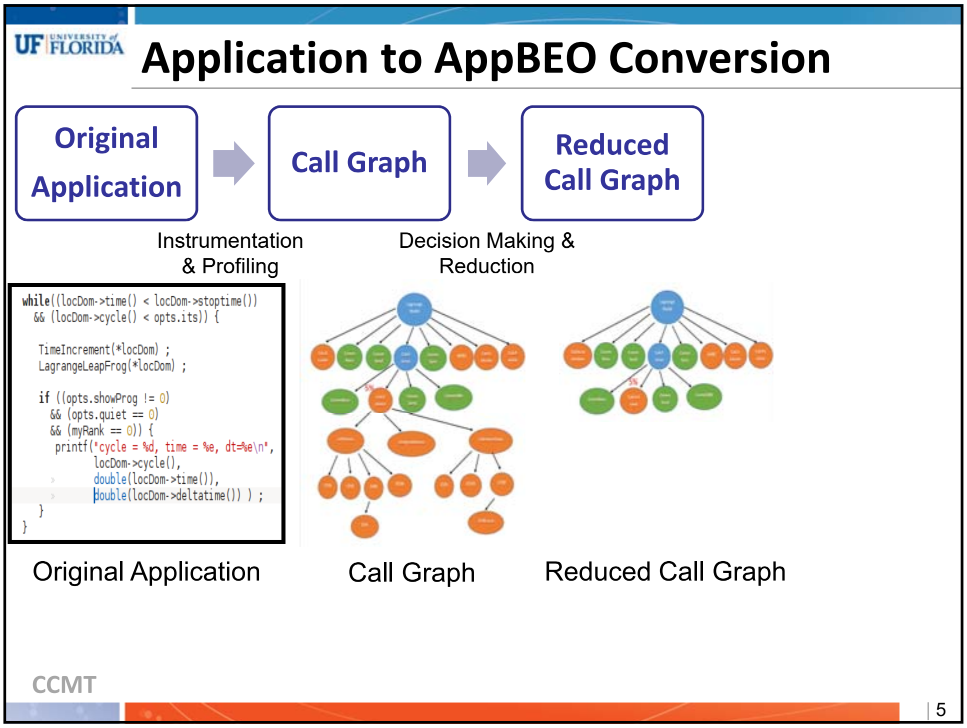
Instrumentation & Profiling

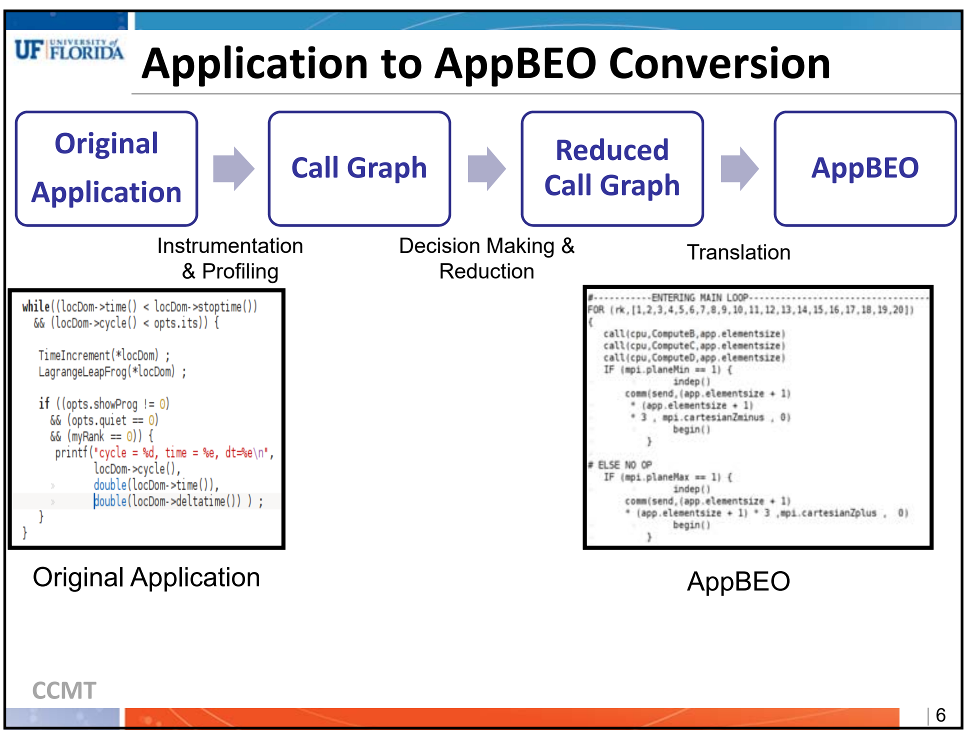
```
while((locDom->time() < locDom->stoptime())
  && (locDom->cycle() < opts.its)) {

  TimeIncrement(*locDom) ;
  LagrangeLeapFrog(*locDom) ;

  if ((opts.showProg != 0)
    && (opts.quiet == 0)
    && (myRank == 0)) {
    printf("cycle = %d, time = %e, dt=%e\n",
          locDom->cycle(),
          double(locDom->time()),
          double(locDom->deltatime()) ) ;
  }
}
```

Original Application   Call Graph

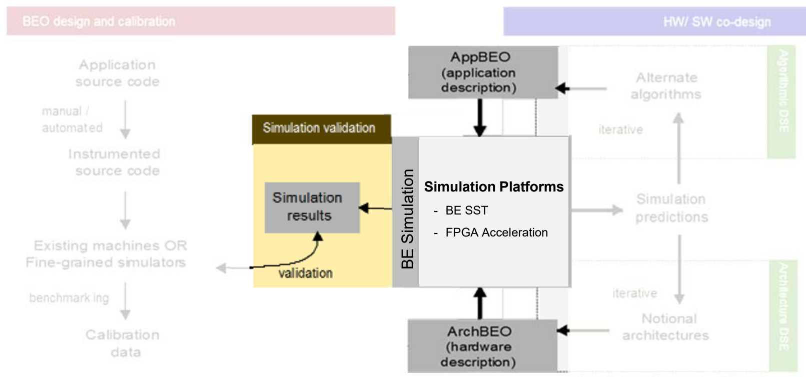
Computation Block
Communication Block
Computation and Communication Block

CCMT | 4


UF UNIVERSITY of FLORIDA
Application to AppBEO Conversion
Original Application
Call Graph
Reduced Call Graph
Instrumentation & Profiling
Decision Making & Reduction
while((locDom->time() < locDom->stoptime())
&& (locDom->cycle() < opts.its)) {
TimeIncrement(*locDom) ;
LagrangeLeapFrog(*locDom) ;
if ((opts.showProg != 0)
&& (opts.quiet == 0)
&& (myRank == 0)) {
printf("cycle = %d, time = %e, dt=%e\n",
locDom->cycle(),
double(locDom->time()),
double(locDom->deltatime())) ;
}
}
Original Application
Call Graph
Reduced Call Graph
CCMT
5


UF UNIVERSITY of FLORIDA
Application to AppBEO Conversion
Original Application
Call Graph
Reduced Call Graph
AppBEO
Instrumentation & Profiling
Decision Making & Reduction
Translation
while((locDom->time() < locDom->stoptime())
&& (locDom->cycle() < opts.its)) {
TimeIncrement(*locDom) ;
LagrangeLeapFrog(*locDom) ;
if ((opts.showProg != 0)
&& (opts.quiet == 0)
&& (myRank == 0)) {
printf("cycle = %d, time = %e, dt=%e\n",
locDom->cycle(),
double(locDom->time()),
double(locDom->deltatime())) ;
}
}
#----------ENTERING MAIN LOOP----------
FOR (rk,[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20])
{
call(cpu,ComputeB,app.elementsize)
call(cpu,ComputeC,app.elementsize)
call(cpu,ComputeD,app.elementsize)
IF (mpi.planeMin == 1) {
indep()
comm(send,(app.elementsize + 1)
* (app.elementsize + 1)
* 3 , mpi.cartesianZminus , 0)
begin()
}
# ELSE NO OP
IF (mpi.planeMax == 1) {
indep()
comm(send,(app.elementsize + 1)
* (app.elementsize + 1) * 3 ,mpi.cartesianZplus , 0)
begin()
}
Original Application
AppBEO
CCMT
6

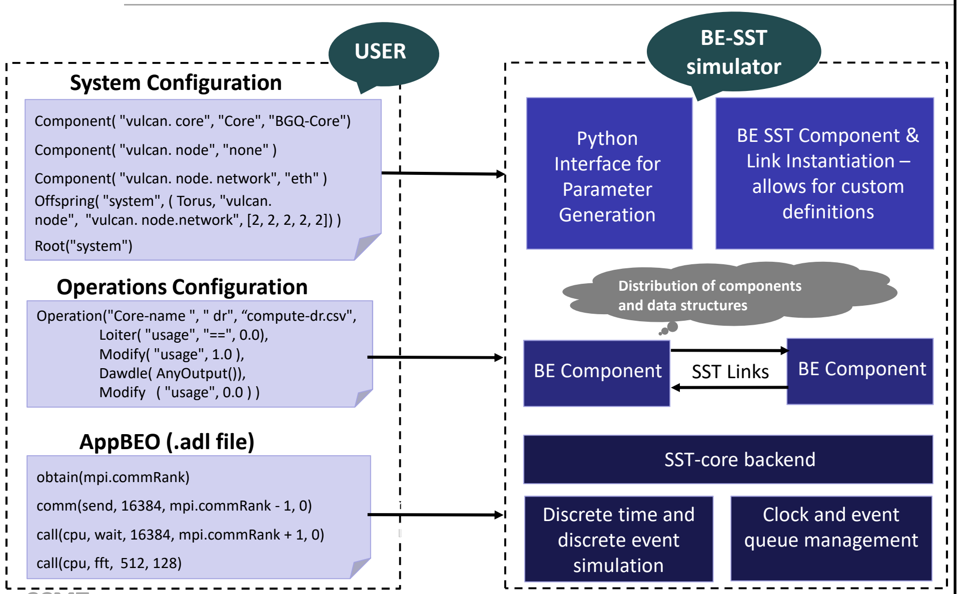
BE-SST Simulator



- SW Platform: **BE-SST** created by extending Sandia's SST
- HW Platform: Novo-G FPGA-based cluster

CCMT | 7

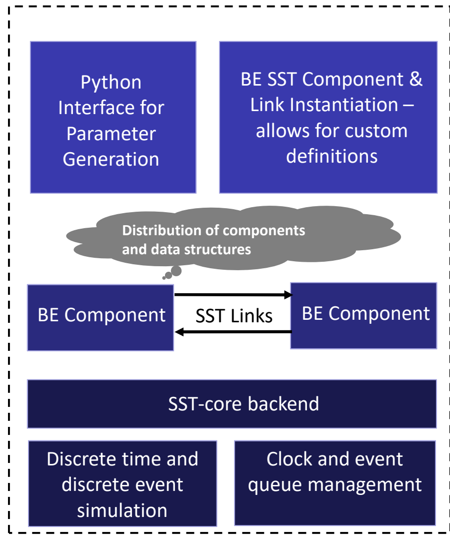
BE-SST v1.0 Features



CCMT | 8

BE-SST model enhancements

- Communication model enhancements:
  - Dynamic generation of network routes provide better performance scaling and parallel performance
  - Overhead modeling at both sender/receiver endpoints
- Interpolation API allows easy switching between different interpolation methods for better accuracy
- Improved scalability and robustness of software product
  - Reduced storage by deleting handled events from queues
  - Switched from static routes to runtime routing



CCMT | 9

Interpolation Interface

- Interpolation interface allows easy switching between different interpolation methods for better accuracy
  - Available schemes: linear, polynomial, lagrangian, kriging

1. Define default in operation configuration file

```
Operation( "BGQ-core", "computeConv", "vulcan-compute-conv.csv", "linear", "elementSize",
        Loiter( "usage", "==", 0.0),
        Modify( "usage", 1.0 ),
        Dawdle( Outputs(0) ),
        Modify( "usage", 0.0 ) )
Operation( "BGQ-core", "prepareFaces", "vulcan-compute-face-proper.csv", "polynomial-4", "elementSize", "elementsPerProcess",
        Loiter( "usage", "==", 0.0),
        Modify( "usage", 1.0 ),
        Dawdle( Outputs(0) ),
        Modify( "usage", 0.0 ) )
```

2. Overwrite with the AppBEO instructions

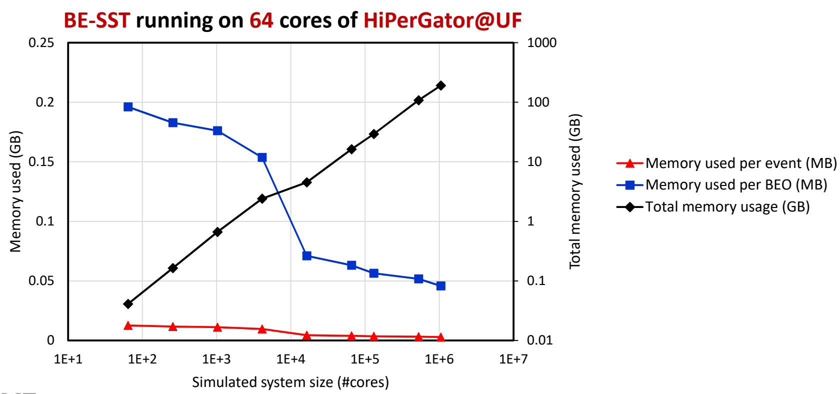
```
call(cpu, computeConv, lagrange, app.elementSize)
call(cpu, prepareFaces, linear, app.elementSize, app.elementsPerProcess)
```

CCMT | 10

UF UNIVERSITY of FLORIDA

Memory Footprint of BE-SST

- BE simulations of different size systems with 3d-mesh topology running CMT-bone-BE
- Observations:
  - Memory used per BEO decreases because each core of HiPerGator is simulating more than one BEO
  - Memory usage has considerably reduced from previous version of BE-SST

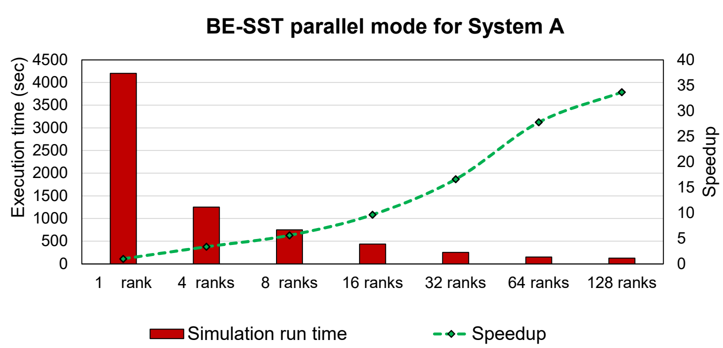


CCMT

UF UNIVERSITY of FLORIDA

Results – Parallel Performance of BE-SST

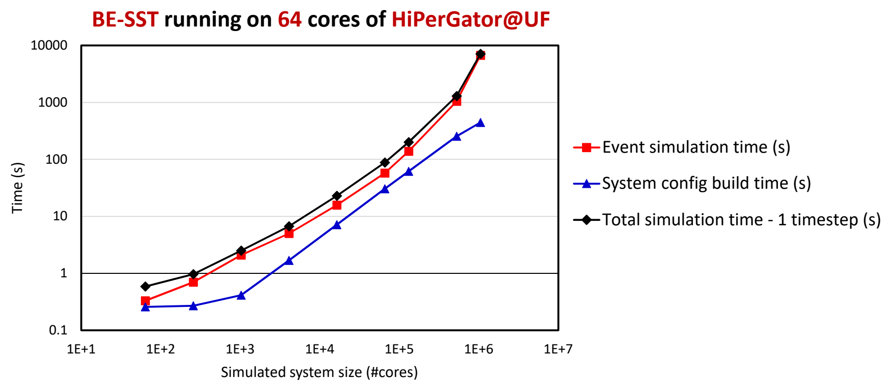
- BE-SST demonstrates good performance scalability
  - Simulation of System A with 131,072 cores executing CMT-bone-BE
  - 385,025 network links in 3D Mesh topology [64, 64, 32]
  - Peak speedup of 33.7x was achieved on 128 MPI ranks



CCMT

Parallel Performance of BE-SST

- BE simulations of 3d-mesh system running CMT-bone-BE
  - Simulations of machines with upto a million cores
  - With dynamic routing
- System config build time is not a bottleneck for large simulations

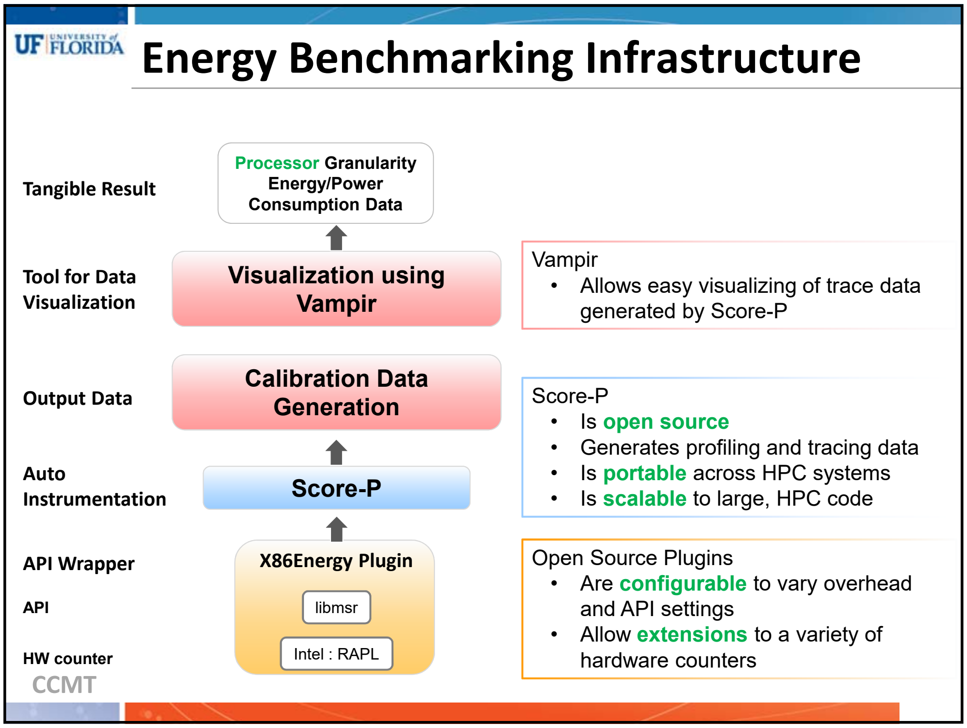


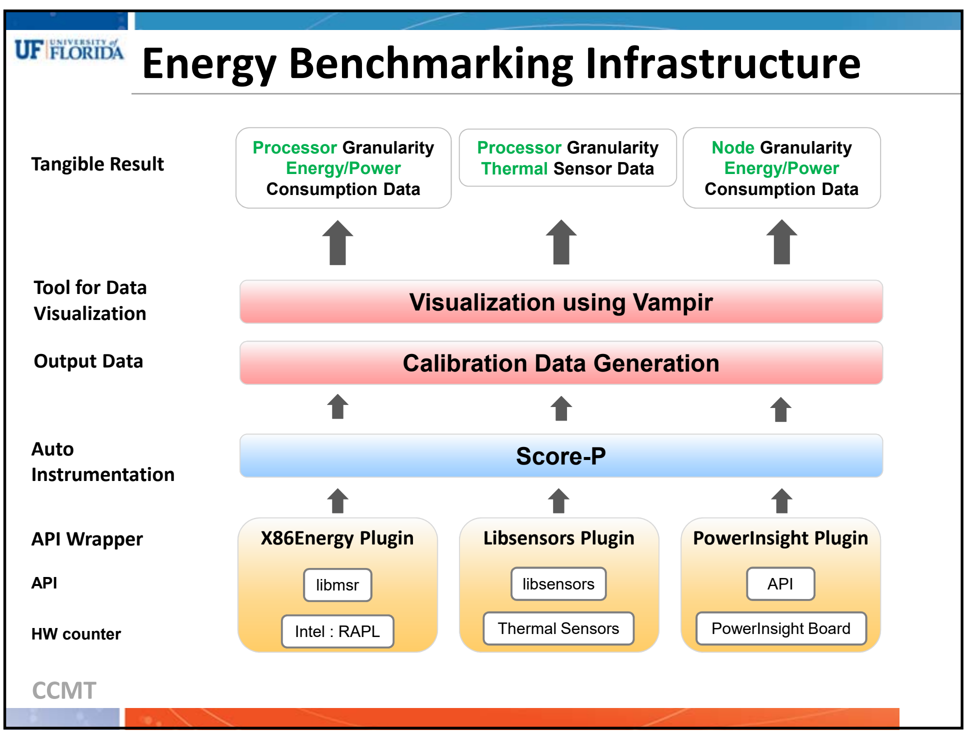
CCMT

13

CCMT

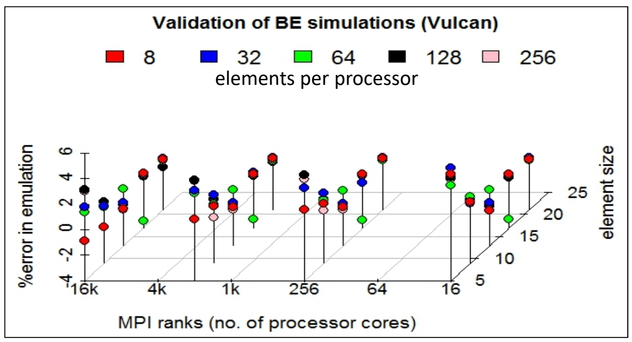
Do you have any questions?


Energy Benchmarking Infrastructure
Tangible Result
Processor Granularity Energy/Power Consumption Data
Tool for Data Visualization
Visualization using Vampir
Vampir
• Allows easy visualizing of trace data generated by Score-P
Output Data
Calibration Data Generation
Score-P
• Is open source
• Generates profiling and tracing data
• Is portable across HPC systems
• Is scalable to large, HPC code
Auto Instrumentation
Score-P
API Wrapper
X86Energy Plugin
API
libmsr
HW counter
Intel : RAPL
Open Source Plugins
• Are configurable to vary overhead and API settings
• Allow extensions to a variety of hardware counters
CCMT


Energy Benchmarking Infrastructure
Tangible Result
Processor Granularity Energy/Power Consumption Data
Processor Granularity Thermal Sensor Data
Node Granularity Energy/Power Consumption Data
Tool for Data Visualization
Visualization using Vampir
Output Data
Calibration Data Generation
Auto Instrumentation
Score-P
API Wrapper
X86Energy Plugin
Libsensors Plugin
PowerInsight Plugin
API
libmsr
libsensors
API
HW counter
Intel : RAPL
Thermal Sensors
PowerInsight Board
CCMT

Validation of BE simulation on Vulcan

- Design of experiment
  - Element size (ES) = 5,9,13,17,21
  - Elements per processor (EPP) = 8,32,64,128,256
  - Number of processors (NP) = 16,256,2048,16384
- Calibration data obtained through in situ benchmarking
- Observation: % error in emulation varies between -3% to 5%

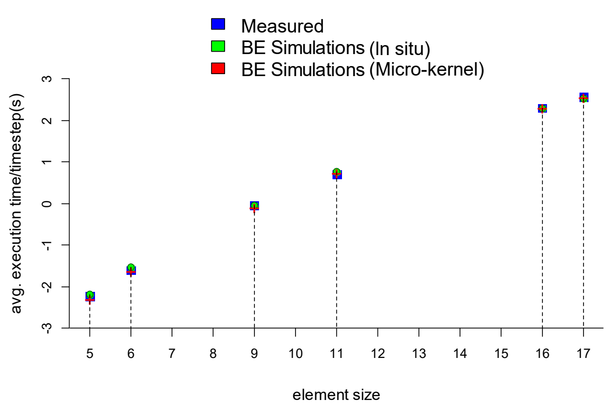


CCMT

17

In situ vs. Micro-kernel Benchmarking

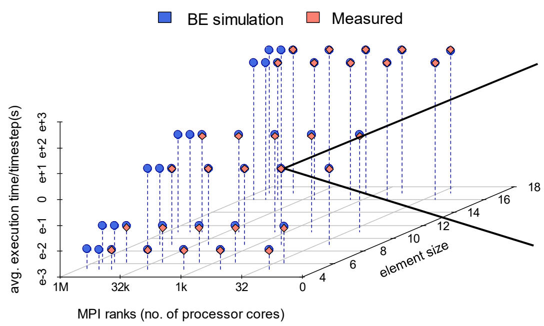
- Micro-kernel benchmarking for calibration to support notional exploration
  - E.g., no source available
- Simulation of 5d torus system with 2048 cores
  - Error is comparable with in situ benchmarking methods for this initial study

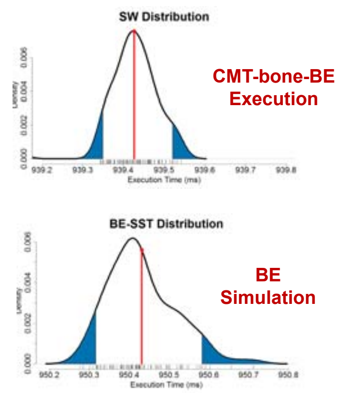


CCMT

18

BE Simulations of CMT-bone-BE on Vulcan





- Average % error between CMT-bone-BE simulation and execution time is 4% and max. error is 9%

CCMT

CCMT

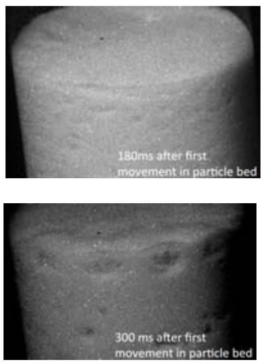
Particle Laden Simulations in CMT-nek

David Zwick

PhD student

Motivation & Setup

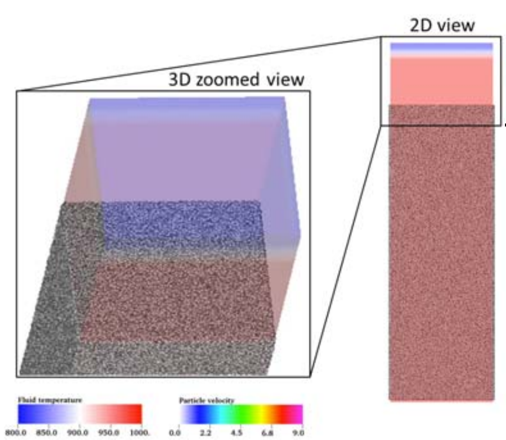
- While not fully resolved, Eulerian-Lagrangian simulations can simulate a realistic number (billions) of particles
- ASU experiment has $O(10^9)$ particles



CCMT

21

One-Way Coupled Results (ASU)



Largest one-way simulations CMT-nek:

Year	Ranks	Elements	Grid pts.	Particles
3	8k	60k	20 million	50 million
4	131k	131k	250 million	20 billion
4	131k	524k	905 million	905 million

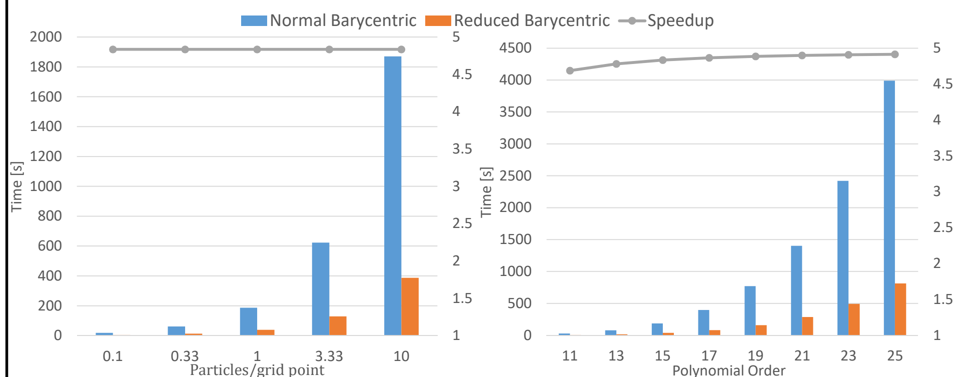
CCMT

22

UF UNIVERSITY of FLORIDA

Collaboration with BE Team

- Full spectral interpolation is costly (normal barycentric algorithm)
- Reduced barycentric interpolation is cheaper
  - 5x speedup on average over full barycentric approach
  - 8x speedup for higher polynomial orders



Ranks	Elements/rank	Polynomial
4096	64	15

Ranks	Elements/rank	Particles/grid pt.
4096	64	1

*Times are total time per time step

CCMT

23

UF UNIVERSITY of FLORIDA

Governing Equations

Fluid Momentum

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot \mathbf{h} = \mathbf{f}$$

Particle Momentum

$$M_p \frac{d\mathbf{V}}{dt} = \mathbf{F}_{hyd}$$

Particle Forces

$$\mathbf{F}_{hyd} = \mathbf{F}_{qs} + \mathbf{F}_{un} + \mathbf{F}_{iu}$$

$$\mathbf{F}_{qs} = \frac{Re_p M_p}{24\tau_p}(\mathbf{u} - \mathbf{V})C_D$$

$$\mathbf{F}_{un} = M_g \frac{D\mathbf{u}}{Dt}$$

$$\mathbf{F}_{iu} = \frac{1}{2} V_p \left(\rho_g \frac{D\mathbf{u}}{Dt} - \frac{d(\rho_g \mathbf{V})}{dt} \right) C_M$$

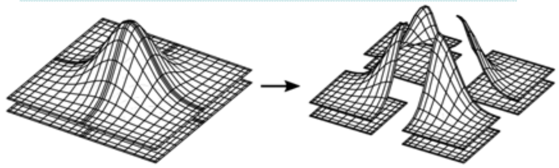
CCMT

24

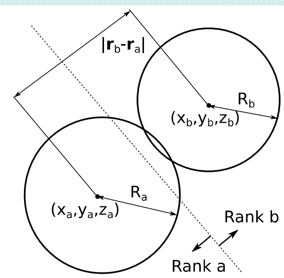
Computational Hurdles in Coupling

Each particle's influence extends beyond its element/rank

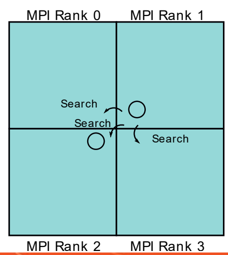
Back coupling with gas for two-way coupling



Particle-particle interaction



Nearest-neighbor search



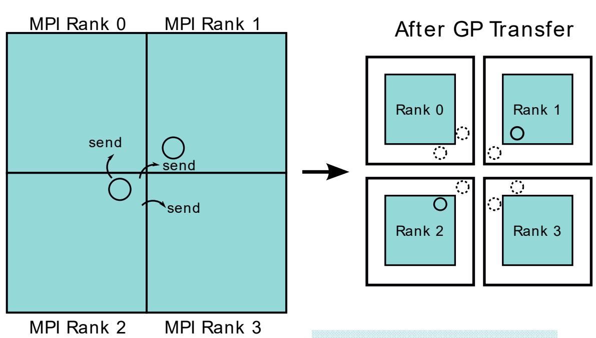
CCMT

25

Ghost Particle Algorithm

Idea:

- If a particle is near a MPI rank edge, it will create a copy of itself called a ghost particle
- Sending perspective rather than receiving perspective



GP steps:

1. Create GP
2. Send GP

CCMT

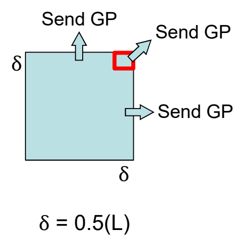
26

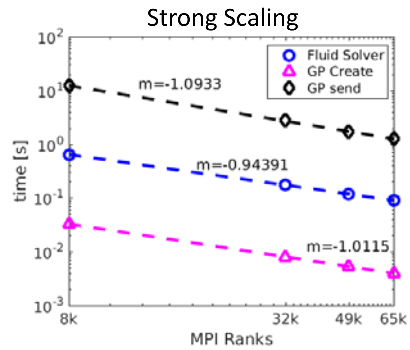
Scaling of Ghost Particle Algorithm

Scalability of algorithm:

- We consider a test case of an expansion wave over a bed of particles
- Particles are randomly distributed and advected by the wave

MPI Ranks	Elements	Polynomial	Grid points	Particles	Time steps
8k–65k	432,000	5	54 million	27 million	2000





*Times are total time per time step per MPI rank

CCMT

Back Coupling with Gas

- Accomplished through the use of a filtering kernel g_m for mollification
- Gaussian kernel is used for g_m
- For example, to transfer particle property A to the grid as a:

$$a(\mathbf{x}) = \sum_{i=1}^{N_p} A(\mathbf{X}^{(i)}) g_m(|\mathbf{x} - \mathbf{X}^{(i)}|) V_p^{(i)}$$

A	a
1	particle volume fraction
$\mathbf{F}_{hyd}/V_p^{(i)}$	particle hydrodynamic force

Fluid Momentum

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot \mathbf{h} = \mathbf{f}$$

For an element in weak form

$$\int_{V_\Omega} v_k f_k dV_\Omega \rightarrow v_k^T B f_k$$

CCMT

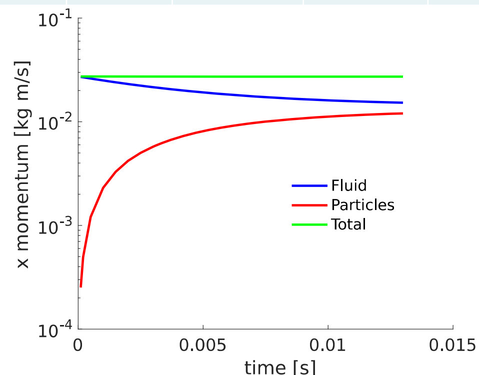
UF UNIVERSITY of FLORIDA

Momentum Conservation

Particles are placed at rest in a uniform flow and subjected to Stokes drag

- Particles speed up and fluid slows down
- Total momentum remains constant

Domain	Elements	Polynomial	Particles	MPI Ranks
x,y,z = [-1,1]	64	11	200,000	64

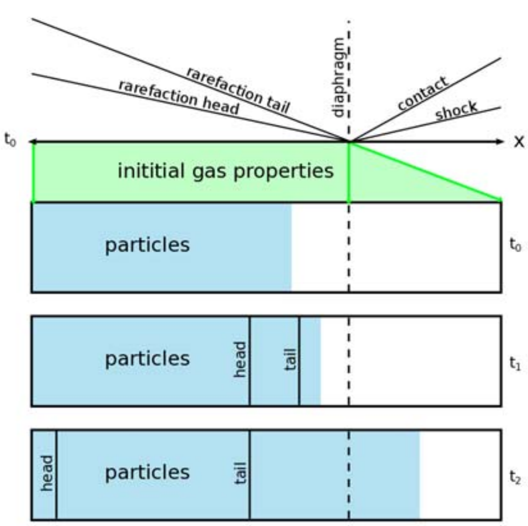


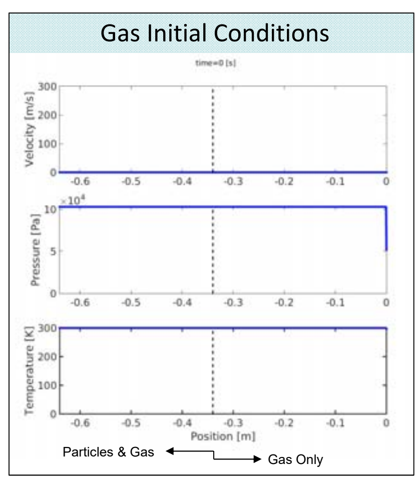
CCMT

29

UF UNIVERSITY of FLORIDA

Application to ASU Experiment





CCMT

30

UF UNIVERSITY of FLORIDA

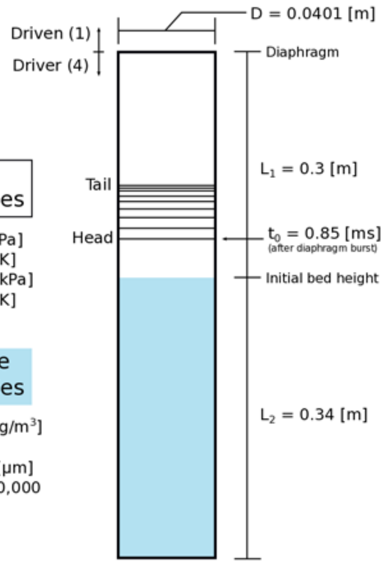
Simulation Setup

fluid properties

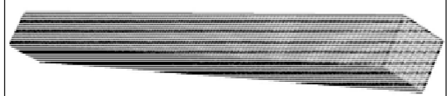
P_1 = 10 [kPa]
T_1 = 294 [K]
P_4 = 101 [kPa]
T_4 = 294 [K]

particle properties

ρ_p = 25 [kg/m^3]
Φ_p = 60%
D_p = 100 [μm]
N_p = 8,000,000



Fluid Grid Setup



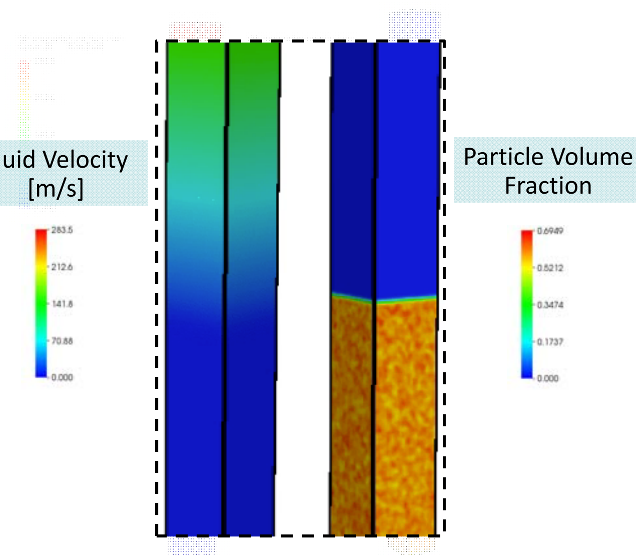
Elements	Grid points	MPI Ranks	SPL
8,192	1,000,000	8,192	76

CCMT

31

UF UNIVERSITY of FLORIDA

Simulation Results

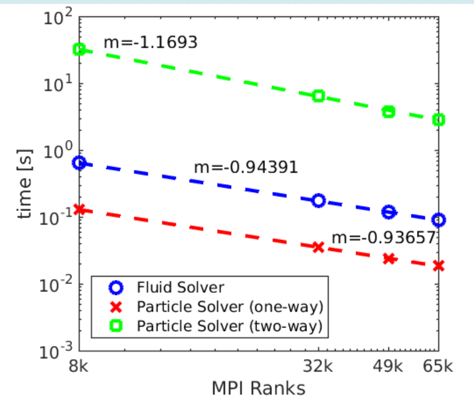


CCMT

32

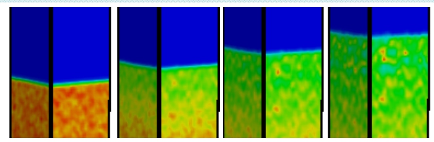
Simulation Results

Strong Scaling



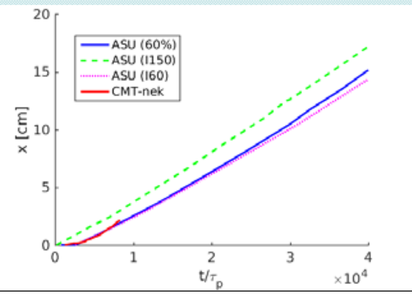
*Times are total time per time step per MPI rank for the same setup as the ghost particle scaling

Void Formation



t = 850.0 [μs] t = 876.3 [μs] t = 884.8 [μs] t = 892.4 [μs]

Bed Height



CCMT

Future work

- Larger simulations on Mira
- Improved force modeling
  - PIEP
  - Volume filtering correction
  - Effects on validation metrics
- Load balancing
  - Conference paper in preparation with Computer Science group
- UQ and V&V

CCMT

CCMT

Do you have any questions?

UF UNIVERSITY of FLORIDA

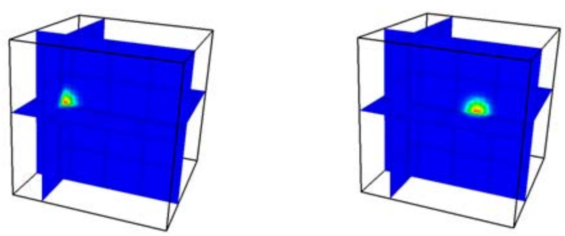
NNSA

UF UNIVERSITY of FLORIDA

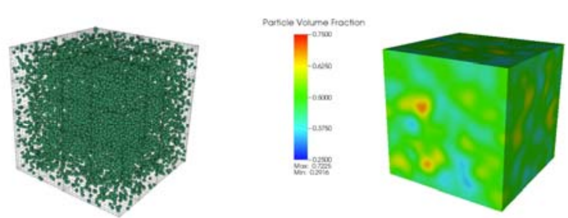
Filtering Examples

- Using ghost particle approach, a mollification kernel is used to spread properties to the Eulerian grid

Force from a single particle



Volume from a cloud of particles



CCMT

36

CCMT

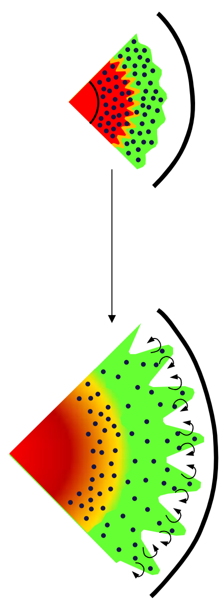
Surrogate Modeling of the JWL Equation of State

Frederick Ouellet
PhD Student

UF UNIVERSITY of FLORIDA

NNSA

The Mixed-Cell Problem with Real Gas E.O.S.



- A short time after detonation, there is pure explosive product (red) and pure air (green)
- At this point, use real gas equations in red cells and ideal gas equations in green cells
- At later time, there are still red and green cells but also a mixing layer shown by the orange cells
- Here, both equations of state must be used

CCMT

38

Standard Approaches for Mixed Cells

Method of Tackling	Pros	Cons
Mass fraction weighted averaging	• Algebraic equations • Small data storage	• Does not place species equilibrium requirement
Tabulated data	• Interpolation is only operation • Can place species equilibrium requirement	• Must store the tabulated data • Interpolation error
Polynomial regressions/curve fits	• Algebraic equations • Small data storage • Can place species equilibrium requirement	• Curve fit errors
Iterative schemes	• Small data storage • Can place species equilibrium requirement	• Computationally expensive

- Rocflu uses a Broyden's method iterative solver to enforce pressure and temperature equilibrium between air and the explosive products

CCMT

Looking at the Big Picture

- What is the iterative scheme really doing?

Inputs:
ρ_m, e_m, Y_{exp}

↓

Outputs:
1) Common pressure/temperature for both EoS
2) Internal energy attributed to the air and the products
3) Density attributed to the air and the products

- If we compute the common pressures and temperatures outside of the code, a curve fit can be applied

CCMT

Models Implemented in Rocflu

The Jones-Wilkins-Lee (JWL) equations of state are used to predict the pressures of high energy substances and are:

$$P_{JWL}\,(\rho, e) = A(1 - \frac{\omega}{R_1 V})e^{-R_1 V} + B(1 - \frac{\omega}{R_2 V})e^{-R_2 V} + \omega\rho e$$

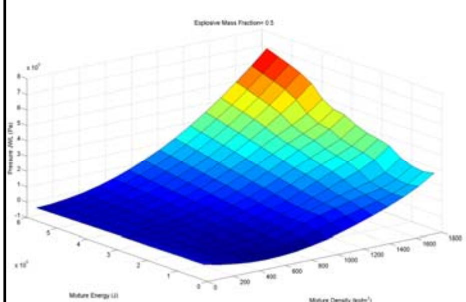
$$T_{JWL}(P, \rho) = (\frac{1}{\rho\omega C_v})(P - Ae^{-R_1 V} - Be^{-R_2 V})$$

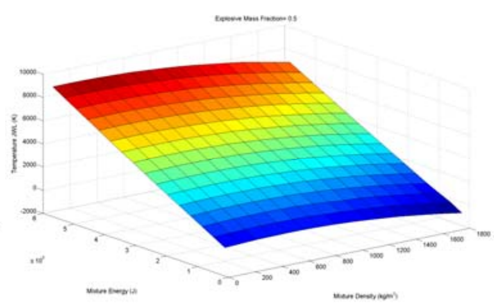
where $V = \frac{\rho_o}{\rho}$ and $\rho_o, A, B, C, R_1, R_2$ and ω are parameters for the substance.

	One Equation Model	Iterative Method	Multi-Fidelity Surrogate
Advantages	- Speed - Algebraic equation	- Accuracy - Problem independent	- Speed - Accounts for species equation
Disadvantages	- Uncertainty and error - JWL + ideal gas case only	- May be slow to converge - Computationally expensive	- Uncertainty - Equation of State specific (problem specific)

CCMT

Surrogate Model - Development

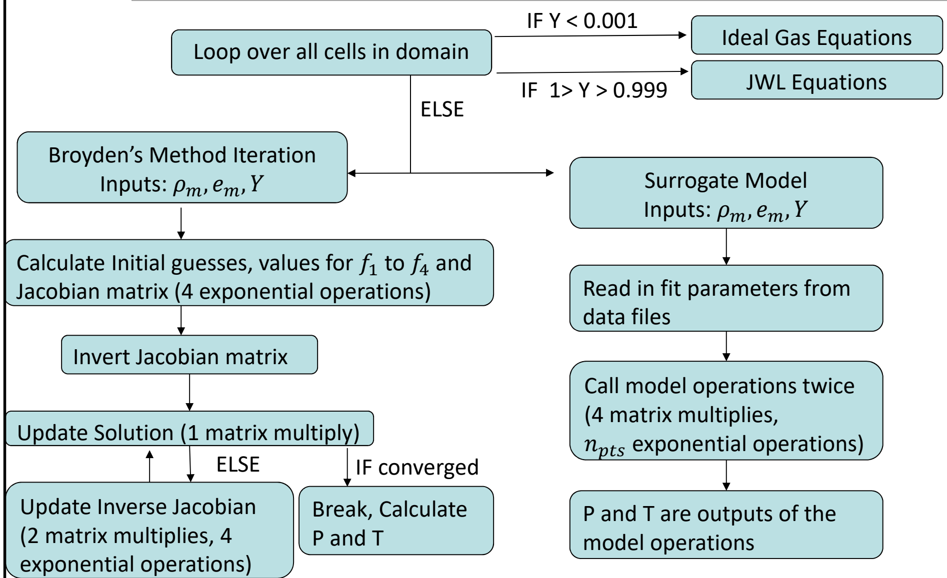




- Models for mixed cell pressure and temperature generated by Kriging with 200 sampling points by UQ team
- Domain Space: $\rho_m \rightarrow (1, 1770)$
  $e_m \rightarrow (1.25e5, 1.5e7)$
  $Y_{exp} \rightarrow (0, 1)$
- Average relative errors:
  - Pressure: 1.298%
  - Temperature: 0.132%

CCMT

Flowchart – Iterative vs. Surrogate

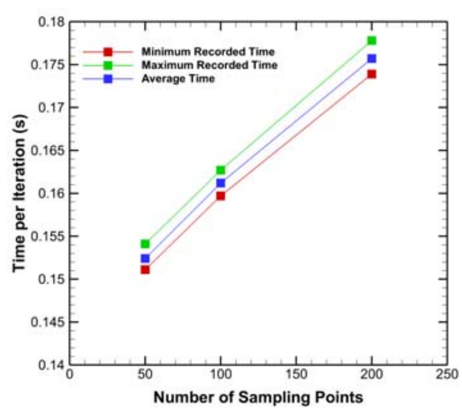


CCMT | 43

Surrogate Model Results – Computation Time

Run	seconds/Time step (Iterative)	seconds/Time Step (Surrogate)	Ratio	Previous Model
1	0.4636	0.1739	2.6662	-
2	0.4314	0.1778	2.4262	-
3	0.4766	0.1754	2.7178	-
Avg.	0.4572	0.1757	2.6024	2.6313

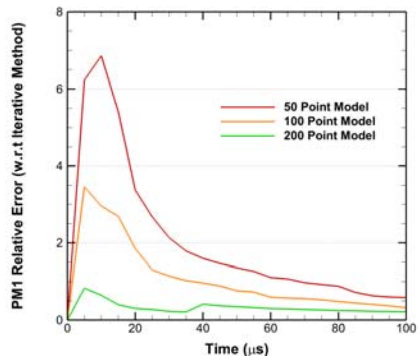
- Average time per iteration in computing 1000 time steps in Rocflu
- Results averaged over three runs
- Speed up ratio for 200 point model is about 2.60



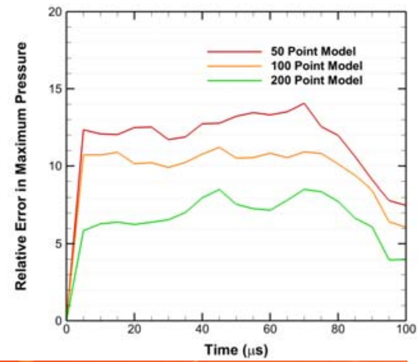
- Preliminary results from study on differing numbers of sampling points
- Adding more points increases computation time

CCMT | 44

Surrogate Model Results – Simulation



- Maximum error in blast wave location for the 200 point model is 0.83%
- Maximum error in peak pressure for 200 point model is 8.51%
- Increasing number of sampling points reduces errors obtained using the surrogate



Simulation Description
- 2D, Quarter-Cylinder grid
- Outer radius = 0.3 m
- 400,000 cells
- Gas-Only

CCMT

45

Future Work

Effect of Sampling Points
- Study of the effect that altering the number of sampling points has to the model error
- Goal is to verify that 200 sampling points hits ideal midpoint of accuracy and computing time

Multiple Species
- Suppose an experiment dealt with a detonation process that contained n species
- If using Rocflu's iterative solver, this means going from a 4×4 system of equations to an $2n \times 2n$ system
- As long as the same number of test points are used to create a new surrogate, its computation time would remain roughly the same

Implementation with Multiphase flows
- Implement the model into simulations of flows with particles.
- Goal is to ensure the surrogate and the iterative scheme have minimal differences in particle trajectories, allowing for confident use in simulating experiments

CCMT

46

CCMT

Do you have any questions?

UF UNIVERSITY of FLORIDA

NNSA

CCMT

Microscale Simulations and Modeling

Yash Mehta

UF UNIVERSITY of FLORIDA

NNSA

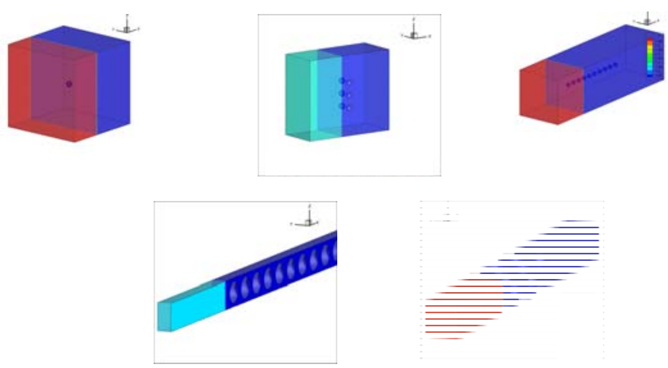
Outline

- Fully resolved simulations of shock interaction with randomly distributed bed of particles :
  - Varying volume fraction
  - Varying shock Mach number
  - 1-D Riemann model for interpreting/analyzing results from numerical simulations
- Pair-wise Interaction Extended Point Particle (PIEP) model

CCMT

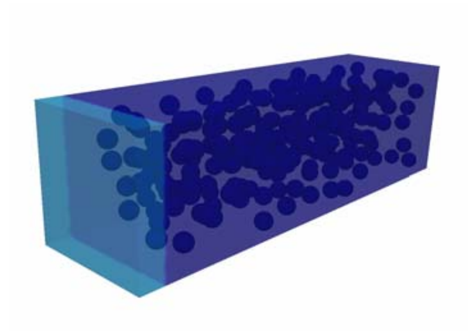
Fully Resolved Simulations

- We have performed multiple fully resolved simulations of shock interaction with particles
- Particle arrangement, particle volume fraction and, shock Mach number was varied
- Different physical mechanisms occurring during shock particle interaction were identified



CCMT

Shock Interaction with Random Bed of Particles

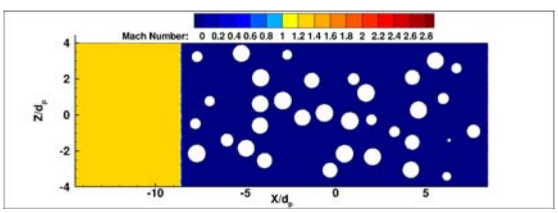


- We have performed multiple simulations by varying the volume fraction and the shock Mach number : $M_s = 1.22, 1.66$ and, 3.00 ; $\varphi = 1.25, 10, 15, 20$ and, 25%
- Number of particles in the computational domain varied from 200 to 500
- Effect of viscosity and particle motion was neglected in these simulations

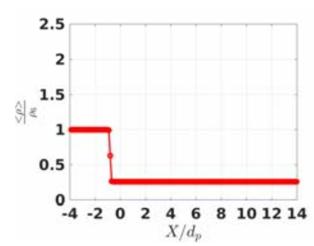
CCMT

51

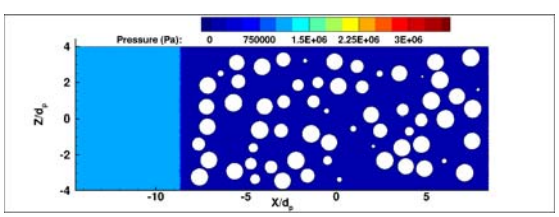
Flow Properties



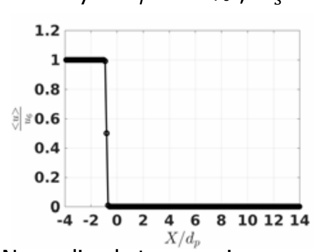
Mach number contour plot for $\varphi = 10\%$; $M_s = 3.0$



Normalized stream-wise averaged density for $\varphi = 10\%$; $M_s = 3.0$



Pressure contour plot for $\varphi = 25\%$; $M_s = 3.0$



Normalized stream-wise averaged velocity for $\varphi = 25\%$; $M_s = 3.0$

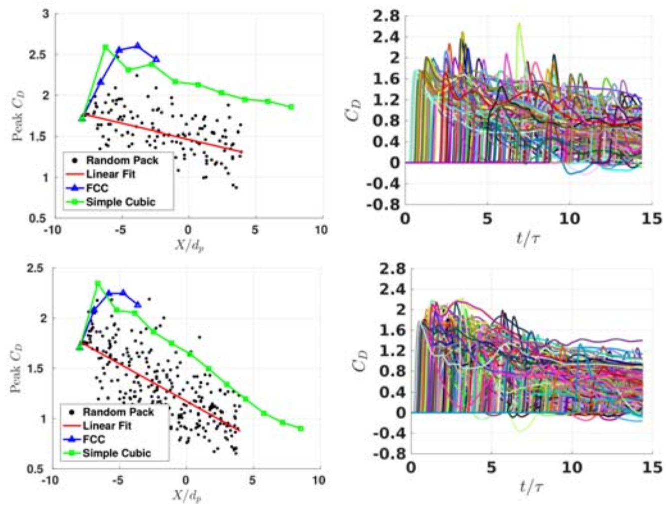
CCMT

52

Forces on particles ; $M_s = 3.0$

$\varphi = 10\%$ $\varphi = 15\%$

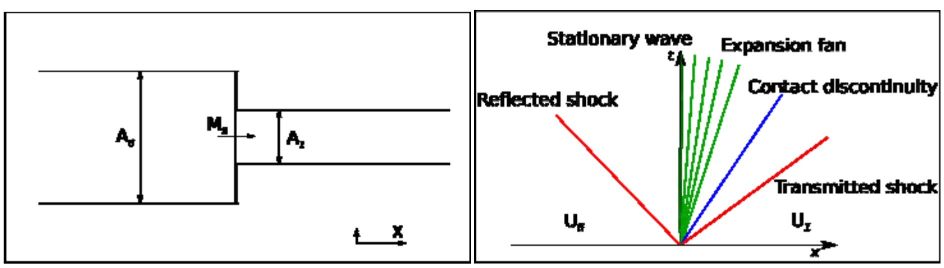
$\varphi = 20\%$ $\varphi = 25\%$



- There is a high variability for the peak drag and a clear downward trend

CCMT

1-D Riemann Model



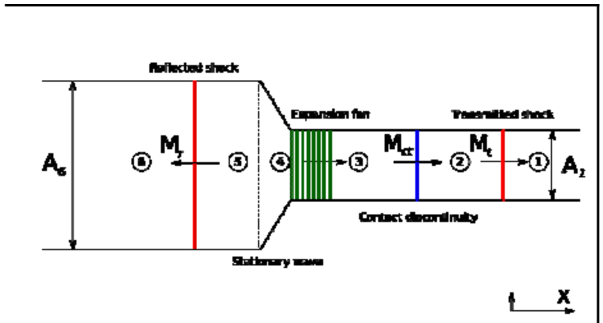
- Shock interaction with bed of particles can be modeled as shock interaction with an area change in a 1-D context
- This is a standard Riemann problem with an additional parameter - area change

$$\frac{\partial \phi}{\partial t} = 0$$

Han etal 2012 "Exact Riemann solutions to compressible Euler equations in ducts with discontinuous cross-section". *Journal of Hyperbolic Differential Equations*, 09(03), pp. 403-449

CCMT

1-D Riemann Model

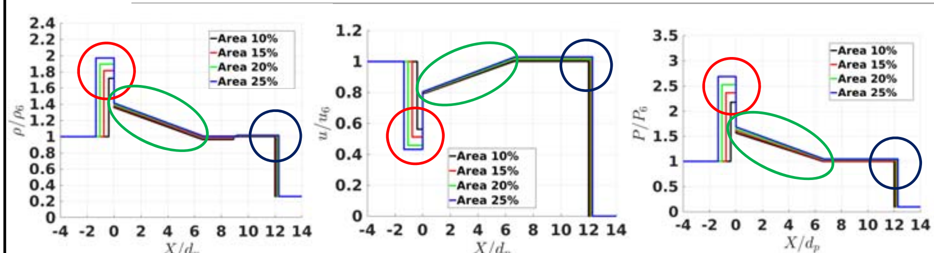


- We obtain a "steady" state solution for different volume fractions
- Reflected shock travels upstream of the area change, there is a resonant expansion fan at the location of area change, contact discontinuity and, transmitted shock travel downstream
- For a given shock Mach number and volume fraction there exists a unique solution

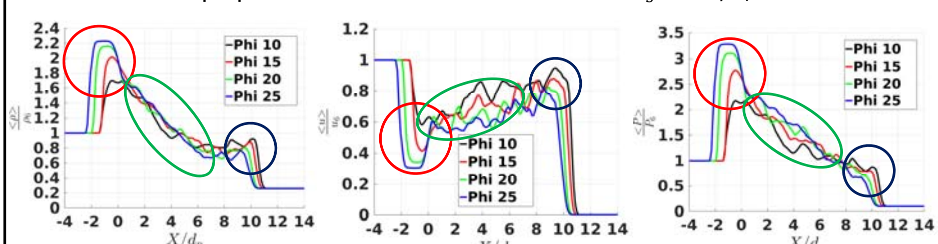
CCMT

55

1-D Riemann Model Vs Numerical Simulation



Flow properties obtained from 1-D model for $M_s = 3.0;\ t/\tau = 12$



Averaged flow properties obtained from numerical simulations for $M_s = 3.0;\ t/\tau = 12$

CCMT

56

1-D Riemann Model Vs Numerical Simulation

φ_1	M_r	M_t	P_5/P_6
10%	1.41	3.00	1.000
15%	1.47	3.03	1.020
20%	1.52	3.05	1.035
25%	1.56	3.07	1.047

Flow properties obtained from 1-D model for $M_s = 3.0$; $t/\tau = 12$

φ_1	M_r	M_t	P_5/P_6
10%	1.41	2.75	0.874
15%	1.59	2.64	0.799
20%	1.67	2.43	0.718
25%	1.72	2.34	0.678

Averaged flow properties obtained from numerical simulations for $M_s = 3.0$; $t/\tau = 12$

CCMT

57

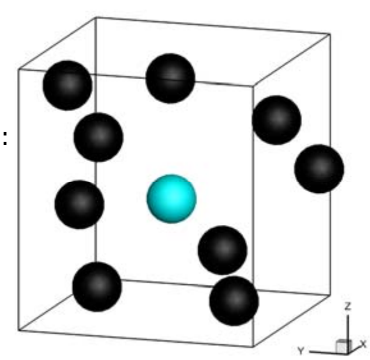
Pairwise Interaction Extended Point Particle Model (PIEP)

- Force on a single sphere is given by the Generalized Faxen's Theorem (F_{flow})
- Force on a single sphere in a bed of particles is :

$$F = F_{flow} + F_{diff}$$

- Force due to presence of other particles (F_{diff}):

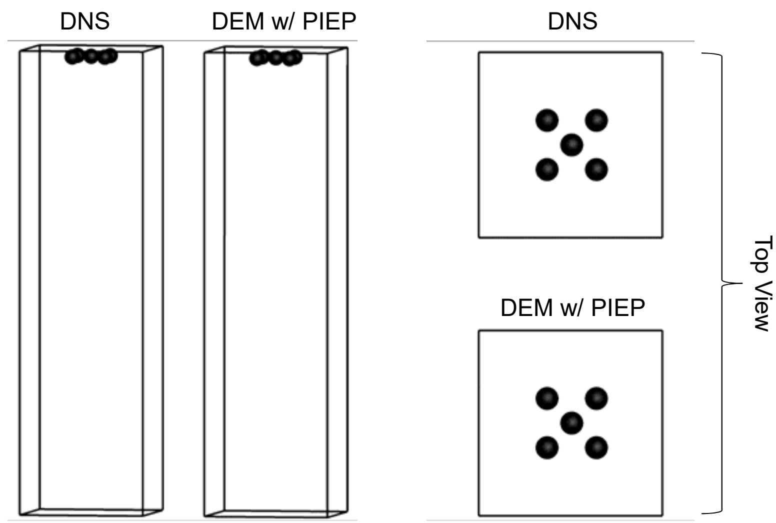
$$F_{diff} = \sum_{j=1}^{N-1} F_{j \to i}$$



CCMT

58

PIEP – Incompressible Flow

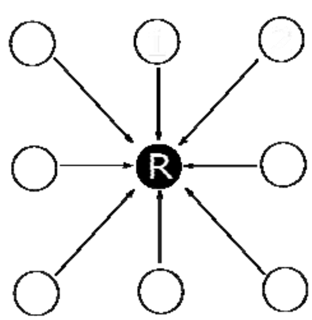


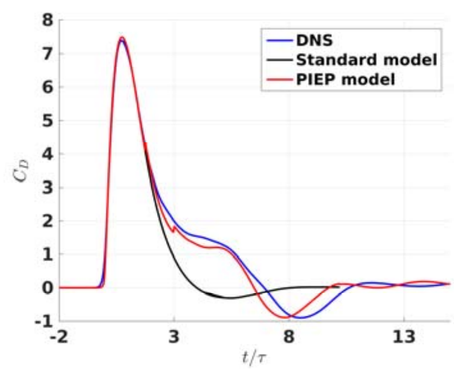
CCMT

59

PIEP – Compressible Flow

- Shock interaction with transverse array of particles





- $F_R = F_{shock} + 4 * F_{1-R} + 4 * F_{2-R}$
- Drag force on the center particle

CCMT

60

UF UNIVERSITY of FLORIDA

Future Work

- Post processing the simulation data
- Viscous simulations for shock interaction with particles
- Developing PIEP model for compressible flows

CCMT

61

CCMT

Do you have any questions?

CCMT

Predicting Execution Times of CMT-nek using Multi-fidelity Surrogate

Yiming Zhang

UF UNIVERSITY of FLORIDA

UF UNIVERSITY of FLORIDA

Accomplishments

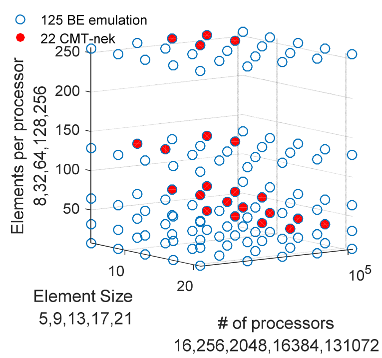
- **Coordinated the multi-fidelity error reduction for CMT codes at large-scale runs**
  - Designed the experiments for comparing CMT-nek (Jason Hackl), CMT-bone (Tania Banerjee) , CMT-bone BE (Aravind Neelakantan) and BE emulation (Aravind Neelakantan)
- **Predicted execution times of CMT-nek using multi-fidelity surrogate**
  - Proposed a multi-fidelity surrogate to fit a few CMT-nek runs together with BE emulations. The multi-fidelity prediction had 4% root-mean-square difference (RMSD) with CMT-nek
  - Submitted to IEEE Cluster 2017 titled as "Using Multi-fidelity Surrogates of Skeleton Apps for Predicting Application Performance"

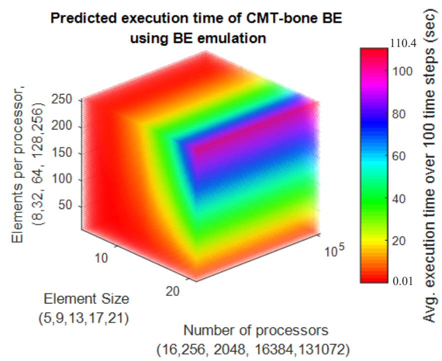
CCMT

64

Design of Experiments at Large-Scale Runs

- Execution times of a typical no-particle test (on Vulcan) using up to 130k processors, 34 million (131072 × 256) elements and 311 billion (21^3 × 131072 × 256) computational grid points
- 125 (5^3) runs for BE emulation & CMT-bone BE, with subsets for CMT-nek & CMT-bone

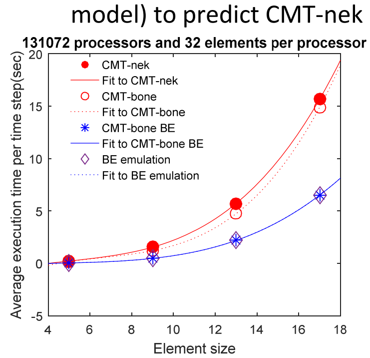


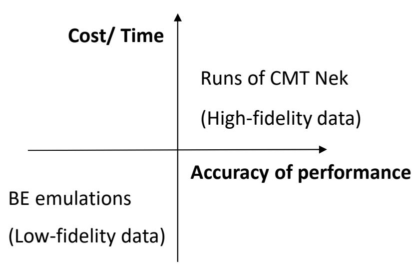


CCMT

CMT codes vs. BE emulation

- Similar trend, but large difference
  - Effective emulation (low-fidelity model) could predict the overall trend well of CMT-nek runs (high-fidelity model)
  - Emulations of CMT-bone BE simplify the calculations, which were much smaller than the execution times of CMT-nek
  - Using multi-fidelity surrogate translates emulations (low-fidelity model) to predict CMT-nek runs (high-fidelity model)





CCMT

Predictions using Multi-fidelity Surrogate

Translate BE emulation against a few CMT-nek runs by polynomial function

Form of translation function

Selected a popular form with a scale factor and a discrepancy surrogate

$$\hat{f}_{nek}(x) = \rho \hat{f}_{BE}(x) + \delta(x)$$

Schemes to determine surrogate parameters

Developed a multi-fidelity surrogate for improved robustness and accuracy

- Bayesian vs. Deterministic
- Spatial distribution vs. Residual error
- Sequential vs. Simultaneous

CCMT

67

Test Plan To Evaluate Multi-fidelity Predictions

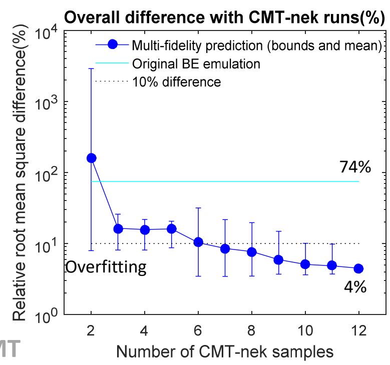
- Selected 10 runs (out of 22) of CMT-nek as validation runs to examine the accuracy of emulation
- The overall difference (root-mean-square) between original BE emulation and CMT-nek was 74% at the 10 validation points
- Adopted 2 to 12 remaining runs of CMT-nek as samples to translate BE emulations using the multi-fidelity surrogate
- For each number of samples, repeated random selection 20 times to account for the effect of sampling plan

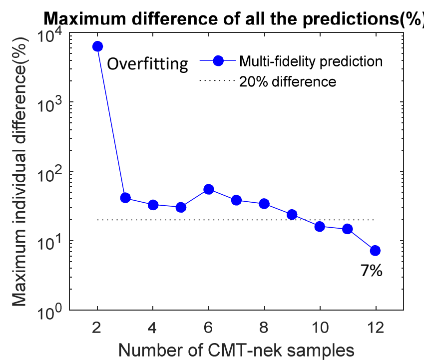
CCMT

68

Multi-fidelity Predictions of CMT-nek Runs

- Overall difference (root-mean-square) is less than 10% with 6 or more CMT-nek samples for correction
- Max error is less than 20% (at the 10 validation runs) with 9 or more CMT-nek samples for correction
- The proposed multi-fidelity surrogate was more accurate than other approaches





CCMT

69

Ongoing Work

- Investigating the capability of corrected BE emulation to extrapolate the execution time of CMT-nek runs towards exascale computing
- Sampling strategy for extrapolating the execution time of CMT-nek runs while incorporating varying cost of runs

CCMT

70

CCMT

Do you have any questions?

UF UNIVERSITY of FLORIDA

NNSA

CCMT

Deeper look into Collapsed Pipeline Approach

Carlo Pascoe

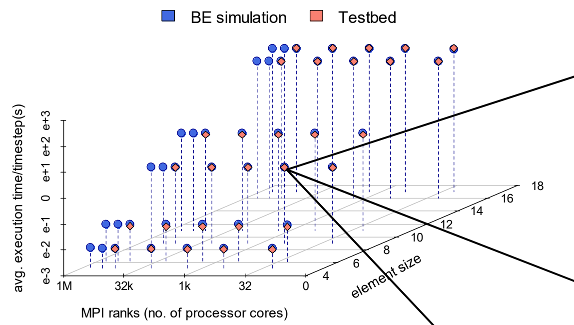
Center for Compressible Multiphase Turbulence (CCMT)

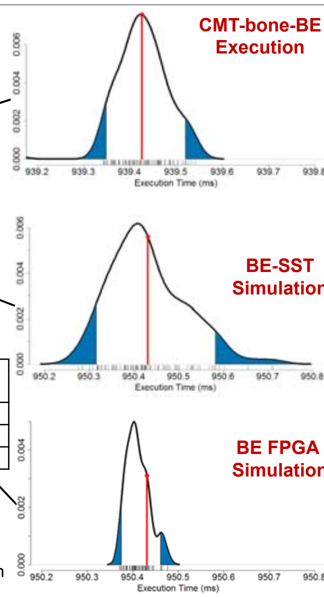
NSF Center for High-Performance Reconfigurable Computing (CHREC)
ECE Department, University of Florida

UF UNIVERSITY of FLORIDA

NNSA

Simulations of CMT-bone-BE on Vulcan





Simulation Method	Num. of MPI Ranks	Num. of Timesteps	Num. of Events	Num. of FPGAs	% Logic Utilization	FPGA Clk Frequency	Latency to First Output	Mega Sims/Sec
BE-SST	128	1	5,952	NA	NA	NA	NA	0.00000152
FEP	128	1	5,952	1	66 %	280 - 320 MHz	66	280 – 320
CP	128	1	5,952	1	2 %	315 - 355 MHz	458	1.68 – 1.99

- BE software and BE FPGA simulations produce similar results
- Fully-expanded approach provides greatest performance at cost of more resources
- Collapsed approach allows for greatly reduced resource utilization and increased scalability at cost of slightly reduced performance

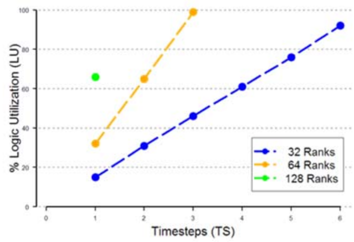
CCMT

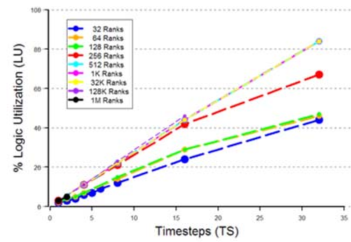
Outline

- Motivation
  - Fully-expanded v Collapsed Pipeline Approach
  - Simple Motivating Example
- Collapsed Pipeline Single-FPGA Performance/Scalability
- Collapsed Pipeline Multi-FPGA Performance/Scalability
- Moving Forward
  - Validation of multi-FPGA predictions
  - Increased performance with Stratix 10
  - Partially-collapsed pipeline approach

CCMT

Fully-Expanded v Collapsed Approach





Fully-Expanded Pipeline

Advantages:

- Superior performance in terms of simulation throughput and latency

Limitations:

- Resources scale linearly with both MPI Ranks and number of timesteps
- Scaling across multiple FPGAs expected to be ineffective when considering exascale simulations

Collapsed Pipeline

Advantages:

- Resources scale linearly with timesteps, but sublinearly with MPI Ranks
- Better scaling on single and multiple FPGAs

Limitations:

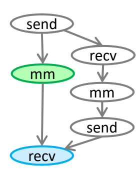
- Lower simulation throughput and longer initial latency, but still more than sufficient for rapid design-space exploration

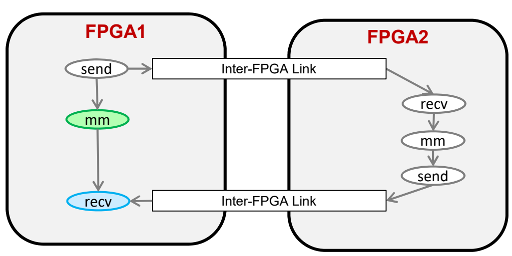
CCMT

75

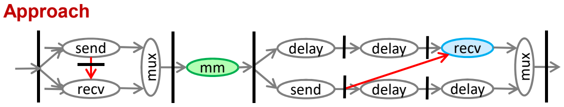
Simple Motivating Example

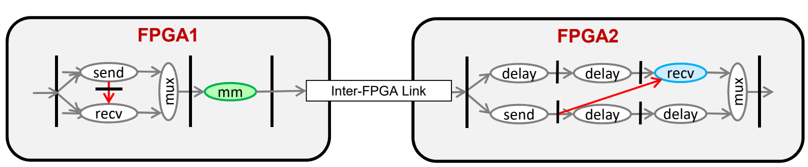
Fully-expanded Approach





Collapsed Approach



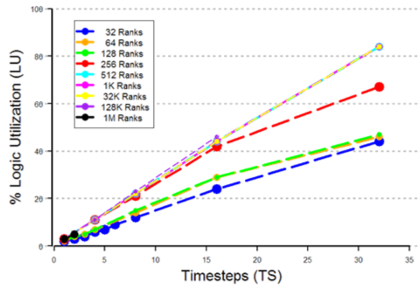


CCMT

76

UF UNIVERSITY of FLORIDA

Collapsed Pipeline Single-FPGA Performance/Scalability



Ranks	TS	% LU	Num. of Events	Lat to First Out (cycles)	Mega Sims /second*	Giga Events /second*
32	32	44	43,008	7,532	10.5	450
64	32	46	92,160	10,252	5.23	482
128	32	47	190,464	10,316	2.62	498
256	32	67	393,216	13,516	1.31	515
512	**32**	**84**	**811,008**	**20,684**	**0.65**	**531**
1K	32	84	1,646,592	21,196	0.33	539
32K	**32**	**84**	**55,443,456**	**241,868**	**1.02E^{-2}**	**567**
128K	16	46	111,673,344	333,932	2.56E^{-3}	285
1M	2	5	112,459,776	1,148,056	3.19E^{-4}	36

As number of ranks increase:
- additional logic per rank approaches zero
- simulation throughput reduced by a factor of "ranks"
- event throughput remains proportional to instantiated event hardware

Pipelines scale linearly with length of simulation, however blockram eventually becomes limiting resource
- e.g., only fit 2 TS for 1 million ranks on Stratix V
- Motivation to explore partially-collapsed pipeline approach

Increased performance expected with Stratix 10
- Collapsed hardware specifically designed to exploit new Stratix 10 architecture

CCMT *Collapsed performance of CMT-Bone-BE with varied MPI ranks and simulation timesteps on a single Stratix V S5GSMD8K1F40C2 @ 335MHz

77

UF UNIVERSITY of FLORIDA

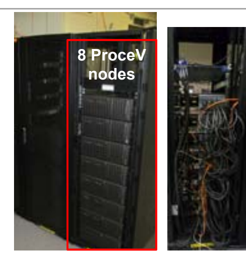
Collapsed Pipeline Multi-FPGA Performance/Scalability

Pipelines scale linearly with the length of simulation, but as a single, unidirectional pipe
- Partitioned easily/predictably across any number of connected FPGAs

Given a desired # of simulated ranks, timesteps, granularity (events per timestep), we can predict performance & # of FPGAs

Main point: FPGAs achieve similar scale as BE-SST, but orders of magnitude faster
- 1M ranks, 300+ of sims/second

Ranks	TS	FPGAs	Num. of Events	Lat to First Out (cycles)*	Mega Sims /second*	Giga Events /second*
1K	32	1	1,646,592	21,196	0.33	539
1K	64	2	3,293,184	41,439	0.33	1,077
1K	128	4	6,586,368	81,925	0.33	2,155
32K	32	1	55,443,456	241,868	1.02E^{-2}	567
32K	64	2	110,886,912	451,039	1.02E^{-2}	1,134
32K	128	4	221,773,824	869,381	1.02E^{-2}	2,267
128K	32	2	223,346,688	536,863	2.56E^{-3}	571
128K	64	4	446,693,376	942,725	2.56E^{-3}	1,142
128K	128	8	893,386,752	1,754,449	2.56E^{-3}	2,283
1M	**32**	**16**	**1,799,356,416**	**2,641,321**	**3.19E^{-4}**	**575**
1M	**64**	**32**	**3,598,712,832**	**4,234,137**	**3.19E^{-4}**	**1,150**
1M	**128**	**64**	**7,197,425,664**	**7,414,540**	**3.19E^{-4}**	**2,299**



Novo-G#
- 64 GiDEL ProceV (Stratix V)
- 4x4x2 3D torus or 5D hypercube
- 6 Rx-Tx links per FPGA
- Measured 32 Gbps per link
- 150 ns latency across links

- Require 64bits/cycle between FPGAs
- 335MHz, 21.4 Gbps, 51 additional lat
- 450MHz, 28.8 Gbps, 68 additional lat
- 500 MHz before link BW saturation

* See poster from Carlo Pascoe

CCMT *Predicted collapsed performance of CMT-Bone-BE with varied MPI ranks and simulation timesteps on multiple Stratix V S5GSMD8K1F40C2 @ 335MHz

78

Moving Forward

- Validation of multi-FPGA predictions
- Increased performance with Stratix 10
- Partially-collapsed pipeline approach

Do you have any questions?

CCMT

79

CCMT

Simulation-Motivated Forensic Uncertainty Quantification of Eglin Microscale Experiments

Kyle Hughes
Graduate Student

April 21, 2017

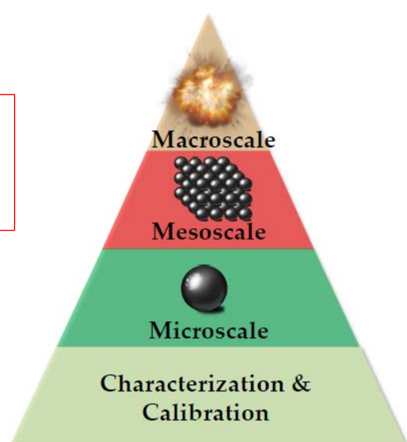
Forensic UQ Perspective

- Significant time lag is present between the microscale experiments and the simulations
  - Not unusual in validation exercises
- As time grows, it becomes steadily more difficult to decrease the epistemic (or lack of knowledge) uncertainty of the experiments
  - Experimental setup or samples are often unavailable for analysis and must be recreated
- Uncertainty may be recognized and reduced through **forensic uncertainty quantification** (UQ) by an independent investigator
- Forensic UQ of the past experiments can be analogous to performing forensics at a crime scene
  - Document all the experimental details (crime scene)
  - Clarify details with experimentalists (witnesses)
  - Quantify uncertain inputs and prediction metrics (forensics lab)
  - Compare to simulation to identify inconsistencies (culprits)
- **Goal:** Apply forensic UQ to the Eglin microscale explosive experiments

CCMT 81

Validation Hierarchy

- CCMT employs a physics-based validation hierarchy
- Eglin conducts experiments at all three scales of interest to CCMT
- **Microscale** (O(1-10^4) particles)
  - Validation of microscale drag models
  - Explosive tests conducted October 2014 and February 2015
- **Mesoscale** (O(10^6) particles)
  - Examination of instabilities, volume fraction effects
  - Explosive tests conducted November 2015
  - Gas gun tests in planning stages
- **Macroscale** (O(10^9) particles)
  - Validation of full simulation
  - Explosive tests (AFRL blastpad) in planning stages



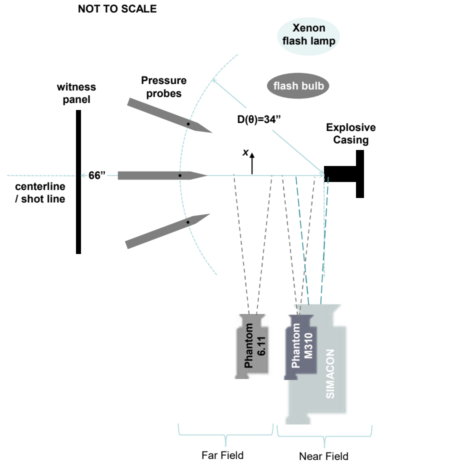
CCMT 82

Microscale Experiments

Side view photograph showing the concave pressure probe array.

Summary of test shots performed in October 2014 and February 2015.

Test #	Particle(s)
Oct14-1	Single (2 mm dia) Tungsten sphere
Oct14-2	Ring of 7 (2 mm dia) Tungsten spheres
Oct14-3	Grid of 16 (2 mm dia) Tungsten spheres
Feb15-1	Single (2mm dia) Tungsten sphere
Feb15-2	Single (2mm dia) Tungsten sphere
Feb15-3	Diamond of 4 (2 mm dia) Tungsten spheres



Overhead schematic of the test set-up.

Sources: Myles Delcambre, Internal CCMT Fall 2014 presentation
Black, Littrell, and Delcambre, Eglin internal written report, 3/7/2015

CCMT

83

Validation and UQ Framework - Microscale

- Experiments
  - Measured Input → Explosive length, Explosive diameter, Explosive density, Casing outer diameter, Casing inner diameter, Particle diameter, Particle density, Particle initial location(s), Ambient pressure, Ambient temperature
    - Measurement Uncertainty
  - Measured Prediction Metrics → Particle trajectory, Contact line location
    - Sampling Uncertainty → Limited # of x-ray exposures, Limited framerates
    - Measurement Uncertainty → User error, Instrument resolution
    - Systematic Uncertainty → Apparatus misalignment, T_0

CCMT

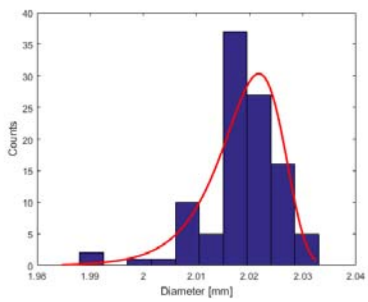
84

UF UNIVERSITY of FLORIDA

Measured Input: Particle Diameter

- Particle diameter was not measured a priori – initial uncertainty was assumed +/- 10% nominal diameter
- Statistical study of tungsten particles was conducted after the experiments to quantify their uncertainty
- Eglin provided 52 tungsten spheres that were randomly measured by two users

Single particle before detonation (Source: Don Littrell, 2/26/2015)



Histograms of particle diameter measurement

- Final data shows uncertainty has been **reduced** to 2-3% of nominal diameter compared to prior belief of 10%
- Smaller range of diameters will also decrease number of runs necessary to construct the surrogate

CCMT 85

UF UNIVERSITY of FLORIDA

Measured Input: Particle Density

- The sample provided by Eglin was additionally used to quantify the uncertainty in the particle density
- Manufacturer cites the particle density as 17 g/cm^3
- All 52 spheres were weighed and then their volume repeatedly measured in a helium gas pycnometer
- Densities are significantly different than those cited by the manufacturer
- Possible causes for this discrepancy are still being investigated
- Energy dispersive x-ray spectrometry may identify the specific alloy

Pycnometer results with summary statistics for the 52 tungsten particles

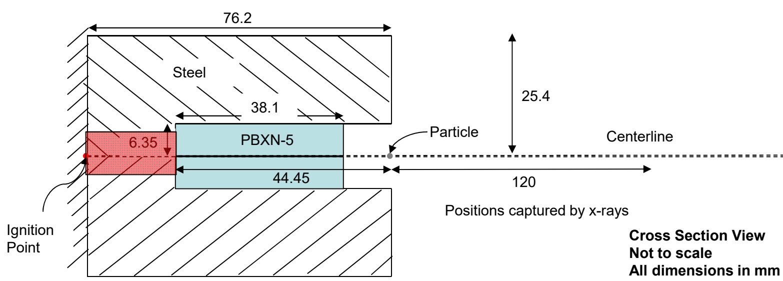
Run	Volume [cm^3]	Density [g/cm^3]
1	0.2401	15.92
2	0.2416	15.82
3	0.2470	15.48
4	0.2546	15.02
5	0.2462	15.53
6	0.2457	15.56
7	0.2429	15.74
8	0.2428	15.75
9	0.2491	15.35
10	0.2489	15.36
11	0.2480	15.42
12	0.2460	15.54
Average	0.2461	15.54
Std Dev	0.00394	0.25

CCMT 86

UF UNIVERSITY of FLORIDA

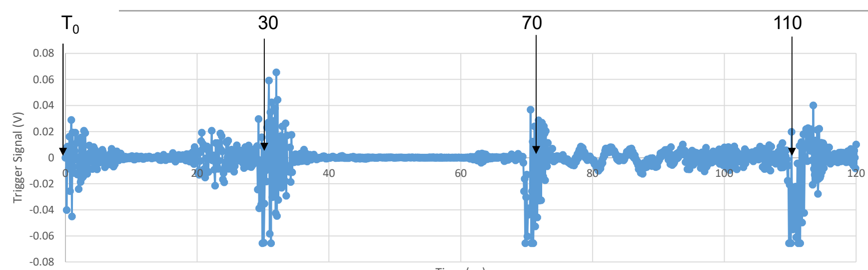
Microscale Simulations T_0

- Large assumptions currently made in the simulations
  - High density quiescent gas is representative of the explosive after reaction
  - Requires time shift to align the experiment and simulation
- To align with experiments need to account for both time of detonator to act and the time for the explosive to burn
  - Manufacturer reported detonator function time: 5.38 ± 0.125 μs
  - Assuming constant detonation velocity through the explosive: 4.31 μs
  - Total delay: ~9.7 μs

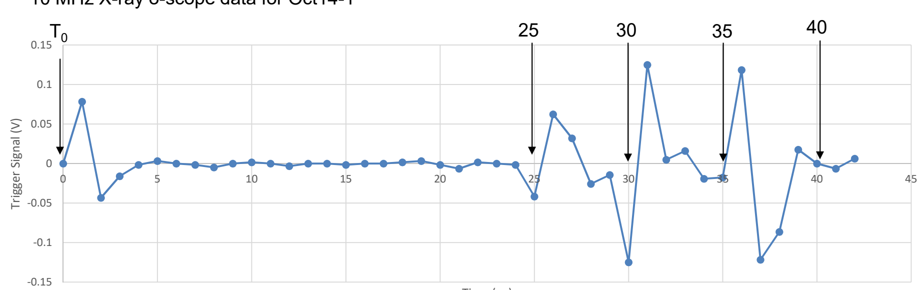


UF UNIVERSITY of FLORIDA

Timing Uncertainty



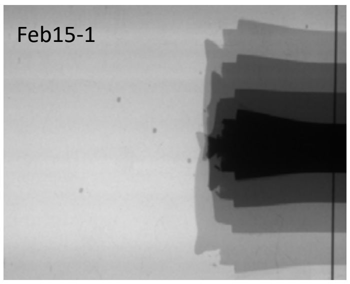
10 MHz X-ray o-scope data for Oct14-1



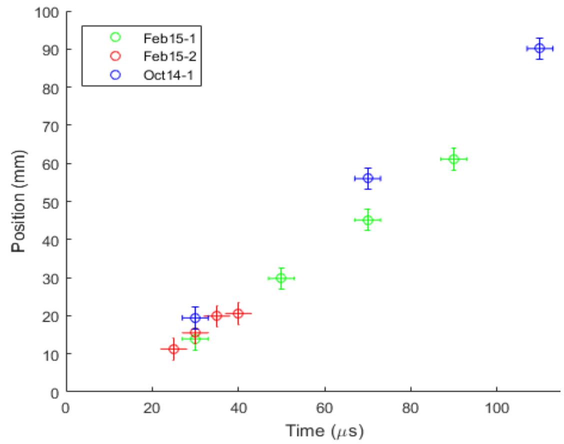
1 MHz X-ray o-scope data for Feb15-2

Metric: Single Particle Position

- Horizontal error bars are determined from oscilloscope records
- Vertical error bars are determined based on spread in data at repeated times



Multiple exposure x-ray of single particle



Single particle position from x-ray data

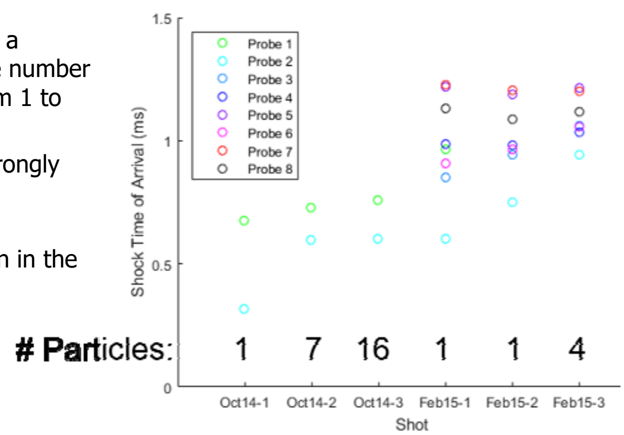
CCMT

89

Shock Time of Arrival

- Initially expected negligible effect on the shock time of arrival based on the small number of particles present
- Shock time of arrivals show a increase in the mean as the number of particles is increased from 1 to multiples
- Evidence that the flow is strongly coupled with particles
- A possible explanation is to examine the volume fraction in the initial plane of the particles:

Number of Particles	Volume Fraction [%]
1	1.65
4	6.6
7	11.6
16	27



CCMT

90

Conclusions and Future Work

Conclusions

- Uncertainty quantification effort for even this "simplified" problem requires extensive back and forth between UQ, simulation, and experimental teams (5+ months)
- Understanding of the simulation assumptions and inputs can help drive forensic investigation
  - Difficulty of determining T_0 resulted in significant investigation into oscilloscope records
- Forensic UQ found a surprise that neither experimentalist or simulationist expected: The small number of particles show significant coupling with the flow

Future Work

- Completion of the uncertainty quantification of remaining microscale uncertain inputs
  - The explosive parameters are especially critical but also challenging
- Construction of surrogates to propagate input uncertainties through simulation
- Comparison of the experimental prediction metrics with simulation results for validation of the drag force models for a single particle

CCMT

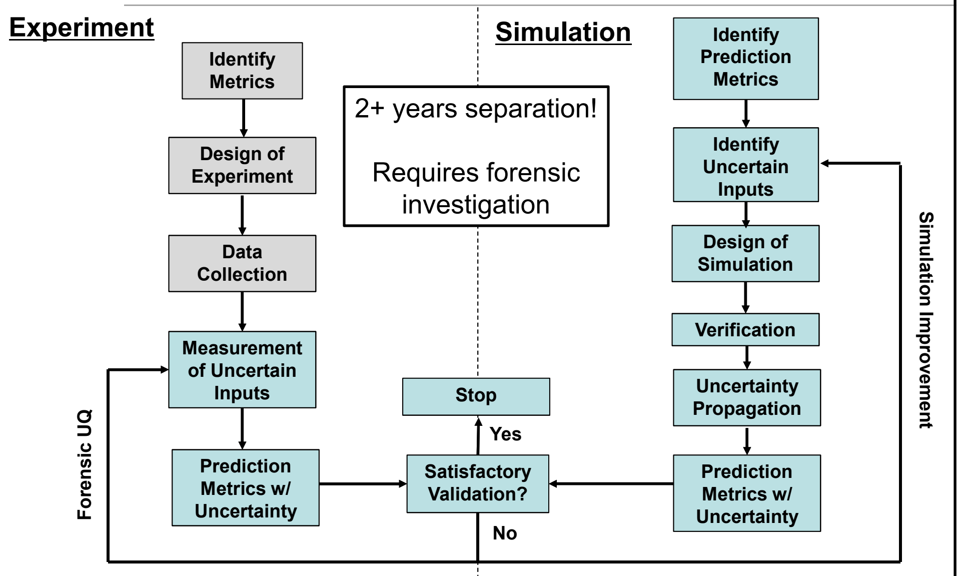
91

CCMT

Do you have any questions?

UF UNIVERSITY of FLORIDA

Forensic UQ Work-Flow



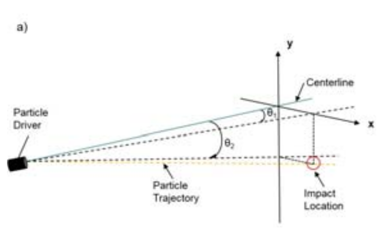
CCMT

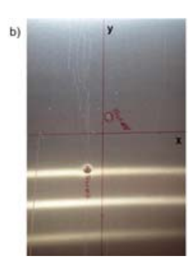
93

UF UNIVERSITY of FLORIDA

Out-of-Plane Motion: Witness Panel

- Witness panel results provides some indication of out-of-plane movement of the particles by providing final position
- Single particle shots show very small angles of the particles
- Repeated particles show significant spread such that two of the particles did not impact the panel while the other two impacted fairly close to the shot line
- The small angles of the single shot results allows elimination of out-of-plane movement as a large possible source of uncertainty





Angles obtained for February 2015 shots

	θ_1 (deg)	θ_2 (deg)
Feb15-1	-0.6494	-1.5718
Feb15-2	0.2392	0.7177
Feb15-3	2.8685	6.1961
	-6.5337	-5.0452

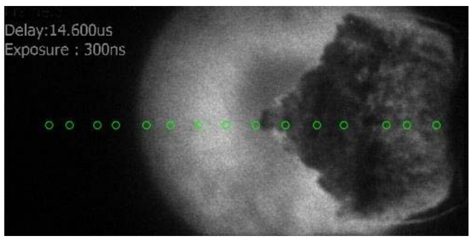
a) Angle conventions used for the witness panel results b) Raw witness panel results

CCMT

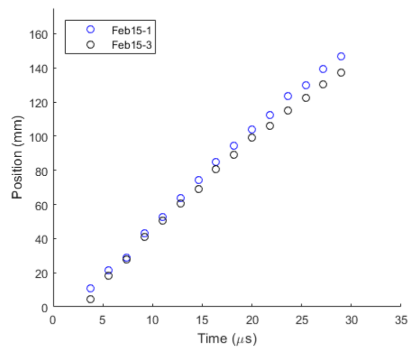
94

Metric: Contact Line

- Obtained from the high speed Simacon images (maximum 16 frames)
- Contact line results are obtained by the selection of the furthest location from explosive casing
- Contact line is only visible in a couple of frames in October 2014 shots (excluded)
- Uncertainty quantification of the results is on-going



Sample Image



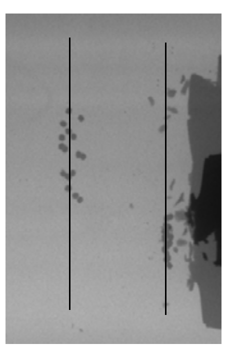
Contact Line Results

CCMT

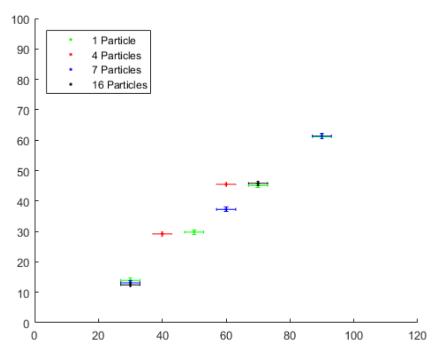
95

Metric: Average Particle Position

- Can be difficult to track each particle for the multiple particle shots
- An average of the particle packet is selected here to begin looking at the multiple particles
- Variance seems reduced by this choice of an average, very similar trends observed with respect to a single particle



Choice of average position for 16 particle shot (Oct14-3)



Average of multiple particles

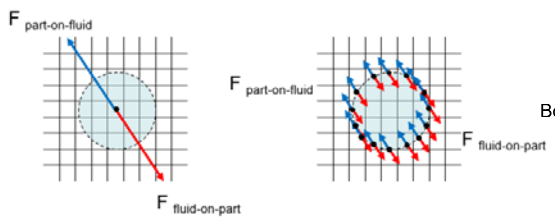
CCMT

96

Simulation Capability

- Clear discussion of simulation capabilities is necessary for design of a validation experiment and forensic UQ
- Current simulation is missing a variety of physics
  - For instance, it does not include effects from particle non-sphericity and deformability
  - Experimentalists did an excellent job in minimizing this effect by using high-density, spherical tungsten particles
- Experimental evidence for significant particle-fluid interaction was surprising for both experimentalists and simulationists
  - Challenging to implement 2-way coupling for such large particles

Point Particle (Current)



Immersed Boundary Method (Future)

CCMT | 97

CCMT

Experimental Studies of Gas-Particle Mixtures Under Sudden Expansion at ASU

Heather Zunino
PhD Student
Dr. Ronald Adrian
Advisor

UF | UNIVERSITY of FLORIDA

Problem Statement and Goals

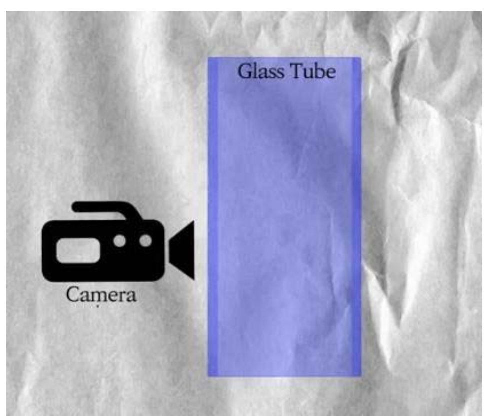
- Experimental multi-phase studies involving compressible flow are complicated
  - Air and solid particles may move separately
  - Particles generate turbulence
- Need for a simple 1D flow experiment that can be used for early validation of the computational codes developed by the PSAAP center.
- Simpler physics involved than the PSAAP capstone experiment
- Reduce the scatter in current data (Chojnicki, et al.)
- Perform experiments on existing shock tube setup
- Determine improvement points and weaknesses
- Design an improved, simple 1D compressible multi-phase flow shock tube experiment
- Examine expansion fan, flow structures, turbulence, and instabilities
- Provide data for early-stage validation of computational codes developed by the PSAAP Center

CCMT 9

UF | UNIVERSITY of FLORIDA

Experiment Description

- 1 meter glass tube
  - Cylindrical footprint
  - Inner diameter: 2”
- Particle bed
- Diaphragm
  - Tape
- High-speed camera
- Measurements
  - Contact line velocity
  - Gas velocity
  - Particle volume concentration
  - Particle interface
- Parameters: particle size and pressure ratio



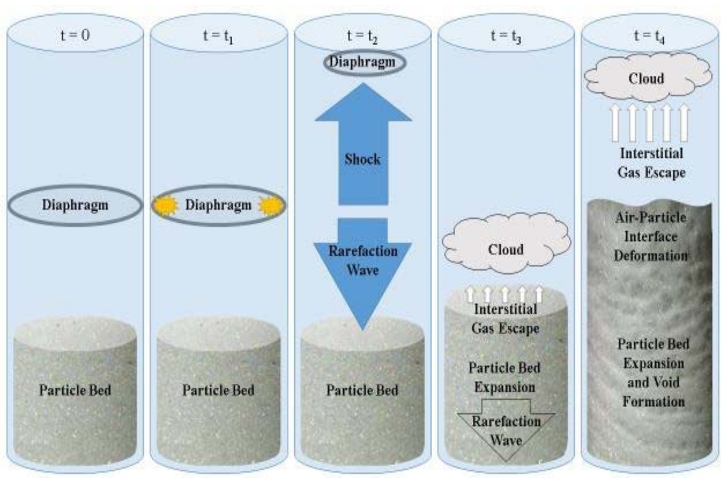
CCMT 1

Experiment Description

CCMT

1

Stages of Experiment

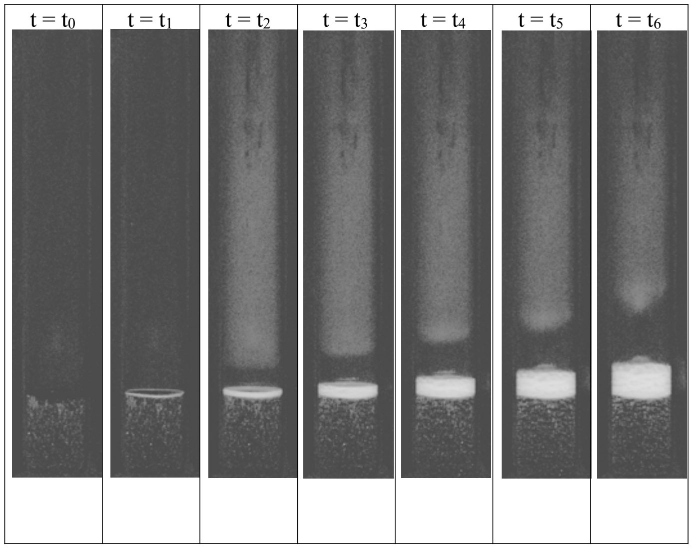


CCMT

1

UF UNIVERSITY of FLORIDA

Cloud Evolution



CCMT

1

UF UNIVERSITY of FLORIDA

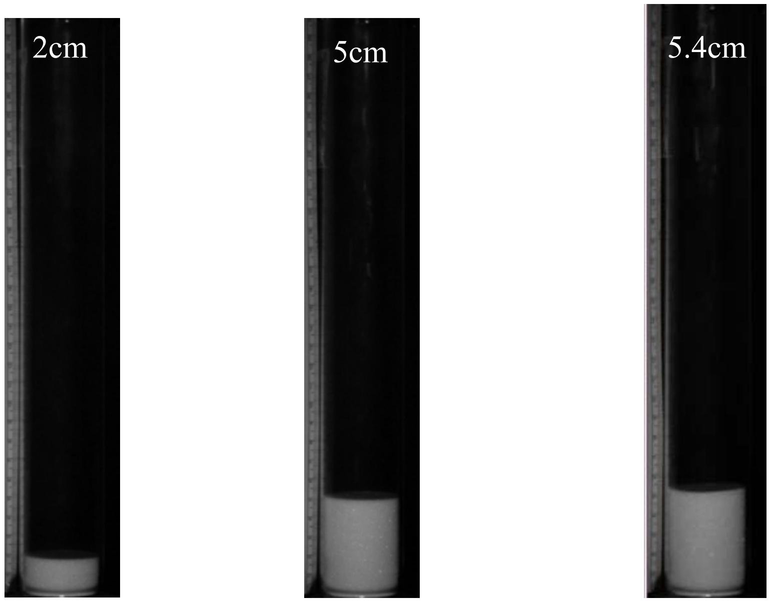
Reflected Shock Wave

CCMT

1

UF UNIVERSITY of FLORIDA

Interface Rise



CCMT

1

UF UNIVERSITY of FLORIDA

Hardware Additions

- 8 pressure sensors
  - 4 locations
  - 90º or 180º apart
- Data Acquisition
  - 8 simultaneous channels
  - Over 150kHz each channel
- Synchronization
  - Camera
  - Pressure sensors
  - Laser light sheet

CCMT

1

Summary

- Cloud evolution
  - Interframe appearance
  - Measure gas velocity escape
  - Gas escapes from edges of particle bed first
- Reflected shock wave
  - Use cloud droplets as gas tracer
  - Particles seem unaffected
- Interface rise
  - Different bed heights
  - Different initial volume fraction
- Hardware additions
  - Upgraded pressure sensors
  - Better temporal resolution
  - Synchronized data acquisition

CCMT

10

CCMT

Do you have any questions?