Accelerating predictive simulation of IC engines with high performance computing (ACS017)

K. Dean Edwards (PI),

Charles E. A. Finney, Wael R. Elwasif National Transportation Research Center Oak Ridge National Laboratory

DOE EERE Sponsors: Gurpreet Singh, Michael Weismiller Vehicle Technologies Office

2018 DOE Vehicle Technologies Office Annual Merit Review

18-21 June 2018 Arlington, VA, USA

This presentation does not contain any proprietary, confidential, or otherwise restricted information.

This work was performed by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy.

ORNL is managed by UT-Battelle for the US Department of Energy



ational Laboratory

Overview

Timeline

- Project start FY2012
- Currently proposed for FY2017-2019

Barriers

- Targets key barriers from USDRIVE roadmap for modeling and accelerating development of advanced engines with improved efficiency and emissions
 - "... develop more robust, computationally efficient models for combustion system design for improved efficiency and reduced CO2 emission. The knowledge base and modeling tools are required to design combustion systems for maximum fuel economy and minimum emissions." 2018 ACEC Roadmap

Budget

- FY2016 \$450k
- FY2017 \$340k
- FY2018 \$400k

Partners

- Leveraging DOE Office of Science ASCR leadership computing resources
 - Multiple ALCC and DD allocation awards
 - Currently totaling ~30 Mhrs on Titan and other OLCF resources
- Multiple collaborative efforts with industry (OEM and ISV) and NLs



- Collaboration with LLNL for application of GPU-enabled Zero-RK chemistry solvers on Titan (ACS012)
- Strong connectivity to broader simulation portfolios at ORNL and DOE
 - Co-Optima simulation team: FT052, FT053, FT055
 - Daimler SuperTruck II team: ACS100
 - Advanced propulsion materials: MAT057
 - Engine knock prediction with The Ohio State University



Overall Relevance and Approach

Supports collaborative efforts with industry (OEMs and ISVs), other NLs, and academia using simulation and HPC to accelerate development of advanced engines with improved efficiency and emissions

• Leverages use of DOE's leadership high-performance computing (HPC) resources

Oak Ridge Leadership Computing Facility (OLCF)



- DOE User Facility funded by the Office of Science through the Advanced Scientific Computing Research (ASCR) program
- Yearly competitive proposal process for resource allocations: DD, ALCC, INCITE*



Titan





ACCEL: OLCF Industrial Partnership Program

- Provides path for industry to collaborate with ORNL researchers on complex problems that require HPC resources
- Open, precompetitive (early TRL) projects (inputs may remain proprietary)
- Expectation that results will be published in open literature



- Supported tasks address technical barriers that are...
 - Of particular interest to the industry partner... and transportation industry as a whole
 - Within scope of DOE-VTO research focus areas
 - Well-suited for the use of DOE's leadership HPC resources
 - Precompetitive (early TRL) with results publishable in open literature
- Multi-year efforts must be coordinated through yearly OLCF project allocations
- Today's supercomputers are tomorrow's workstations

* DD = Director's Discretion (typically 1-5 Mhrs)
 ALCC = ASCR Leadership Computing Challenge (5-20 Mhrs)
 INCITE = Innovative and Novel Computational Impact on Theory and Experiment (20+ Mhrs)



Current Task List and Milestones

Projects / Milestones	Collaborations / Status
Enabling virtual engine design and calibration with HPC and GPUs HPC and GPU-based solvers to enable increased simulation detail and predictive accuracy to advance virtual engine design and calibration	GM 🏈 💥 🖳
 2015 and 2016 ALCC projects using GPUs to enable practical use of highly detailed chemistry in ICE CFD simulations Milestone: Evaluate impact of enabling increased chemistry detail on predictive accuracy and computational requirements 	Completed July 2017 Completed FY2017-Q1
 2017 ALCC project to add further model fidelity Full-cylinder, open-cycle simulations Conjugate heat transfer (CHT) Milestone: Evaluate impact of full-cylinder model and CHT on predictive accuracy and computational requirements 	In progress On track for FY2018-Q4
 Proposed 2018 ALCC project to add LES and further develop CHT 	Submitted Feb 2018
Investigating initiation of autoignition and knock in GDI engines HPC for CFD engine simulations with coupled LES and CHT to study onset of autoignition	
 New task in FY2018 planned as multi-year project DD project for initial development and proof-of-concept 	Kick-off held Dec 2017



Investigating initiation of autoignition and knock in GDI engines





CAK RIDGE National Laboratory

5 ACS017 ORNL – KD Edwards

Detailed modeling of autoignition with HPC – Relevance, Approach

Relevance

- FCA, ORNL, and Convergent Science (CSI) recently launched a new collaborative effort to investigate fundamental factors promoting autoignition and knock in GDI engines
- Using HPC to enable higher fidelity engine simulations than possible with conventional computing resources
 - LES with multiple realizations through flow-field variation to examine impact on knock
 - CHT for spatially and temporally varying thermal boundaries to identify autoignition locations

Approach

- Planned as a multi-year effort using HPC to enable improved detail in simulation of autoignition
 - Include model details and physics not practical with conventional computing resources
 - Systematic approach to adding detail
- Final goal: detailed CFD modeling of autoignition conditions using...
 - LES turbulence modeling
 - High spatial (mesh) and temporal (CFL) resolution to resolve in-cylinder high-speed pressure oscillations
 - Significantly increases computational requirements (~3x or more)
 - Variation of in-cylinder flow field to mimic cyclic variability and its effect on knock
 - Conjugate heat transfer (CHT) modeling to identify hotspots and autoignition locations within cylinder
 - Investigation of effects of multiple spark plugs on engine knock
- OLCF DD allocation of 4 Mhrs on Eos awarded late-Nov 2017, kick-off in Dec 2017



Detailed modeling of autoignition with HPC – *Accomplishments*

Technical Accomplishments

- Finalized baseline model of modern GDI engine in CONVERGE
- Initial simulation efforts underway
 - Spark sweeps performed using RANS turbulence model for baseline
 - Verified that model captures knocking onset relative well



- LES turbulence model has been added
 - Grid-resolution study with LES underway

Ongoing and Future Work*

- Finalize selection of LES base mesh size
- Apply perturbations to flow field in LES cases to study their impact on engine knock
- Finalize CHT model and conduct initial simulations

* Proposed future work is subject to change based on availability of funding and allocation of HPC resources



Enabling virtual engine design and calibration with HPC and GPUs





Virtual design and calibration – *Relevance*

Virtual design and calibration has potential to significantly accelerate engine development



Currently not very practical with conventional computing resources

- Classic speed vs. accuracy trade-off
- Requires simulations at 100s or 1000s of operating conditions
- Very simple models needed for acceptable throughput on conventional computing resources
 - Sector geometry
 - Closed-valve simulations with initial cylinder conditions from 1-D models (e.g., GT-Power)
 - Skeletal kinetic mechanism or simple combustion models
- Negative impact on accuracy
 - Reduced accuracy
 - Applicable over limited range of operating conditions
 - Often need to "tune" model for differing conditions across full operating space



Virtual design and calibration – *Relevance*

Virtual design and calibration has potential to significantly accelerate engine development



HPC and GPU-enabled solvers enable *higher-fidelity* CFD simulations and *faster throughput*

More detailed physics- and chemistry-based submodels

- More predictive, less tuning
- Single model capable of covering full calibration map with little or no tuning
- Increased throughput with parallel simulation
 - Simulate entire engine map in one batch
- GPU-based solvers for computational speed up
 - Example: Sector-model simulations with detailed mechanism (766 species, 6787 reactions)
 - On Titan with Zero-RK solvers and GPUs: complete (IVC to EVO) in ~5 days
 - With conventional solvers on CPU-based clusters: stalled after ~15 days only completing from IVC to +10 °ATDC

Today's supercomputers are tomorrow's workstations

- GPU-based systems increasing available and affordable
- Methods and techniques developed on HPC now will be ready for future commercial systems



Virtual design and calibration – Approach

GM, ORNL, LLNL, and CSI are partnering on multi-year effort to use HPC and GPUs to examine impact of increased simulation detail on predictive accuracy and computational requirements

- Goals:
 - Improved accuracy
 - Increased prediction (less tuning)
 - Increased coverage of operating space without tuning for each case
 - Identify which simplifications are most important to remove
 - <u>Approach applicable to CI and SI</u>

• Systematic approach to add detail to CFD model and assess impacts on accuracy and speed





Virtual design and calibration – Approach

Baseline model overview

- GM 1.6-L midsize Diesel
- Sector-mesh, closed-valve, RANS model in CONVERGE (v2.3)
- GT-Power simulations provide initial cylinder conditions at IVC _
- Skeletal n-heptane mechanism (47 species, 74 reactions), Zel'dovich NOx
- 600+ cases in DoE of engine control parameters over full engine calibration ٠
 - 10 parameters: Speed, SOI & PW for pilot & main, rail P, EGR, wastegate, etc.
 - Includes extreme operating points that stretch capabilities of model
 - *e.g.*, injection sweeps, EGR sweeps, rich operation, etc.

Systematic refinement of model

MechanismRef	Mesh	Other modifications	
Skeletal	Coarse	Baseline model	
Reduced	Coarse		
Reduced	Refined		
Reduced	Refined	Refined spray model; Refined spray-wall interaction	

Colors in table correspond to plots on following slides

Mechanism	Fuel Surrogate	Species	Reactions	Soot
Skeletal ERC n-heptane + Zel'dovich NOx	nC7H16	47	74	Hiroyasu + PSM model
Reduced Chalmers n-heptane + Stanford PAH + GRI NOx	nC7H16	144	900	Hiroyasu

Mesh	Base Grid	Minimum grid with embedding & AMR	Max Cell Count
Coarse	2 mm	0.5 mm	~170k
Refined	1 mm	0.25 mm	~615k

GM 1.6 Liter Midsize Diesel Engine



Speed-load map of 600+ DoE cases





Global correlation shows improved applicability to full speed-load range with model refinement

- Peak cylinder pressure (PCP), CA50: Consistent improvement
- HC: Significant improvement with chemistry and wall-spray refinements
- NOx: Significant improvement with chemistry and grid refinements
- CO and soot: Need further work



13 ACS017 ORNL - KD Edwards

ASME ICEF2017-3631



Significant improvement in HC emissions for individual cases

- Chemistry detail and wall-spray refinements play important roles
- Some cases are under-predicting HC with refined wall treatment
 - May need to further investigate wall-film sub-model



Transitioning from sector model to full-cylinder, open-valve combustion model

- Expected to improve prediction of in-cylinder conditions and charge motion
 - Gas-exchange
 - Production-accurate port and cylinder geometry including <u>positionable swirl valve</u> in intake runner
 - Plume-to-plume spray interaction
- Maintaining previous refinements
 - Mesh: 1.4-mm base grid with 0.35-mm min cell size \rightarrow ~2.5M max cell count (~0.6M for sector)
 - Mechanism: Reduced n-heptane (144 species, 900 reactions)



Full 600-case DoE with this model would not be practical with conventional computing resources



 Initial efforts with full-cylinder model focusing on 20 down-selected points



- Results show continued improvement in CO and soot over the best sector-model results
 - Dramatic improvement in CO accuracy
 - Remaining issues mainly at high load and speed
- Other metrics and emissions remain good or slightly improved



Case #

National Laboratory

RESEARCH CENTER



16 ACS017 ORNL - KD Edwards

Increased oxidation of CO and soot observed with full-geometry model in late cycle

- Secondary "hump" in heat release rate ۲
- Higher turbulence kinetic energy improving • local mixing
- Faster swirl decay late in cycle ٠
- Lower swirl motion and contoured piston top • allow combustion into squish volume earlier and faster





3.0

1.5

1.0

0.5

0.0

0.0

60.0

RESEARCH CENTER

NATIONAL TRANSPORTATION

120.0

-60.0

-180.0

-120.0

Crank Angle (deg. ATDC)

National Laboratory

Virtual design and calibration – *Remainder of FY2018*

Conjugate heat transfer (CHT) for better spatial and temporal thermal boundary conditions

- Thermal BCs are large and important uncertainties for simulations
 - Difficult to measure experimentally
 - Traditionally requires estimates or calibration
- Initial development of CHT model complete (CONVERGE v2.4)
- **Cooling jacket**
 - Steady-state analysis -> not computationally expensive
 - Initial scoping simulations complete
- Combustion chamber with CHT for gas, piston, liner, head
 - Full-cylinder, full-cycle combustion model -> computationally expensive
 - Solve combustion for one cylinder and map thermal BCs to others
 - Super-cycling for CHT solution
 - Initial scoping runs underway at OLCF
 - Severe under-subscription of nodes to avoid out-of-memory

Developing workflow and scripting for full CHT implementation





TRANSPORTATION



Summary of proposed Future Work

Enabling virtual engine design and calibration with HPC and GPUs	
 Finalize workflow and scripting for CHT simulations Complete initial runs with full CHT model Evaluate impact on simulation accuracy 	Remainder of FY2018 Under 2017 ALCC project
 Continued refinement and evaluation of CHT model Add LES turbulence modeling Evaluate impact on simulation accuracy Implement for full DoE 	FY2018-19 * Under 2018 ALCC proposal
Investigating initiation of autoignition and knock in GDI engines	
 Finalize LES base mesh size Apply perturbations to flow field in LES cases to study impact on engine knock 	Remainder of FY2018
Finalize CHT model and conduct initial simulations	FY2018-19 *

* Proposed future work is subject to change based on availability of funding and allocation of HPC resources



Transition from Titan to Summit on the path to exascale

Summit should further enable high-fidelity engine simulations

		SUMMUT	FRØNTIER	
	Titan	Summit	Frontier – 1 st exascale	
	27.1 PF (24.5 GPU + 2.6 CPU)	~5-10x Titan	~5-10x Summit	
	18,688 total nodes 16 (AMD Opteron)	~4600 total nodes 42 (IBM Power9)	?	

6 (NVIDIA Volta)

Memory per node	32 GB (2 GB per core)	512 GB (~12 GB per core) + 96 GB high-band memory	?
Energy usage	9 MW	~15 MW	?
Launch date	2012	2018-2019	Early- to Mid-2020s

• Summit targeted to come online in 2018 (likely 2019 before fully operational)

1 (NVIDIA Tesla)

- Current plan calls for Titan and Summit to co-exist for 1 year before Titan is decommissioned
- Increased memory per core will be <u>significant</u> improvement
 - Current projects forced to under-subscribe nodes (use <16 cores) due to memory limits on Titan
- Scalability and GPU-enabled code will be increasingly important

Current efforts have ORNL well-positioned to take advantage of Summit's capabilities

- Models for ongoing projects are severely taxing Titan's capabilities, especially memory allocation
- But should be well-suited for Summit



Performance

Total Nodes CPUs per node

GPUs per node

Response to reviewer comments

Multiple reviewer comments suggested targeting a smaller portion of the calibration space to focus the scope and reduce the number of required simulations.

Response:

- One goal of this effort is to develop an approach which will provide coverage of the full calibration space with little or no tuning of the model rather than developing a model that only applies to a small subset of conditions. We feel that is best demonstrated by using a large number of simulations that cover a wide range of operating conditions. The current number of simulations (602) are already a small subset of the total number of simulations required for a full engine calibration.
- The need for a large number of individual simulations makes this problem suitable for Titan. Even with the added fidelity of highly detailed mechanisms and large cell counts, the individual simulations only scale to a few dozen cores which could be run on a small GPU-enabled cluster. It is the additional need for rapid throughput of 100s or even 1000s of these simulations which make this effort a suitable application of Titan's massively parallel resources.
- That said, limited allocated hours on Titan under the current 2017 ALCC project and the additional model complexity with fullcylinder geometry and CHT have necessitated a further down-selection to 20 cases during development and proof-of-concept of the CHT model.

Reviewer questioned why full-cylinder models were not used until second year of project.

Response: For the required level of throughput (100s to 1000s of simulations) the appropriate baseline approach with conventional workstations and clusters was determined to be a closed-valve sector model with skeletal kinetic mechanism based on input from our industry partner. Access to Titan requires demonstrated capability to use the GPUs. Thus when prioritizing model refinements to address first, using the Zero-RK GPU-enabled chemistry solvers to enable detailed chemistry had to be the first step in the project.

Reviewer suggested developing "genetic algorithms or Bayesian models" to optimize the calibration process.

Response: OEMs have existing, proprietary tools for optimized engine calibration. While advancing those tools could be a future topic, the current focus is on reducing the time and expense required to produce the data fed into these routines by replacing at least some experiments with simulation. This requires a better balance of the speed/accuracy trade-off that we feel HPC can help provide.

Reviewer suggested use of ANL's SWIFT to manage multiple simulation cases.

Response: While not specifically called out in the presentation, we have our own set of tools and scripts developed at ORNL and OLCF for this purpose which are specifically designed for Titan. ORNL has extensive experience in managing large ensembles of engine simulations dating back to 2012.



Collaborations

Project supports collaborative efforts with industry, NLs, and universities to apply DOE ASCR HPC resources to accelerate development of advanced engines capable of meeting fuel economy and emissions goals

- DOE Office of Energy Efficiency and Renewable Energy
 - Vehicle Technologies Office
- DOE Office of Science
 - Advanced Scientific Computing Research program
 - OLCF User Facility (Titan, Summit)
 - ACCEL Industrial Partnership Program

Enabling virtual engine design and calibration with HPC and GPUs			
General Motors	Provide technical direction, engine model, validation data		
Convergent Science	Support CONVERGE on OLCF machines; assist with model development		
Oak Ridge National Laboratory	Support simulation efforts and HPC resources; provide analysis and direction		
Lawrence Livermore National Laboratory	Support use of Zero-RK GPU-enabled chemistry solvers		
Investigating initiation of autoignition and knock in GDI engines			
Fiat Chrysler Automobiles	Provide technical direction, engine model, validation data		
Convergent Science	Support CONVERGE on OLCF machines; assist with model development		
Oak Ridge National Laboratory	Support simulation efforts and HPC resources; provide analysis and direction		



Remaining challenges and barriers

Applying simulation tools on HPC resources presents unique challenges

- Scalability is vital to maximize benefits of parallel architectures
 - Both for single jobs and ensembles
- Increased model fidelity results in increased memory demands
 - Current HPC systems have limited RAM per core (*e.g.*, 2 GB/core on Titan)
 - Requires undersubscription of nodes to avoid out-of-memory issues -> consumes allocation quickly
- Maximizing benefits of HPC (increased detail, reduced wall time, etc.) must often be balanced with administration rules (queue and scheduling rules, fixed allocations, I/O limitations, etc.)
- Software must continually adapt to evolving hardware technologies
 - Codes need to be increasingly hardware independent
 - OLCF projects must make use of Titan's GPUs
 - Forthcoming DOE-ASCR machines (Summit, Aurora) will further exacerbate this issue
- Software must adapt to scale of HPC environments
 - Memory usage, file I/O, load balancing, restart management, licensing, etc.

Partnering with ISV is crucial for addressing many of these issues



Summary

- <u>Relevance</u>
 - Supports collaborative efforts with industry and others that apply simulation on DOE ASCR HPC resources to accelerate development of advanced engines with improved efficiency and emissions

Approach

- HPC and GPU-enabled solvers enable increased detail in CFD engine simulations compared to conventional resources (workstations, CPU-based clusters)
- Systematic approach to add detail to CFD model and assess impacts on accuracy and speed
 - Detailed chemistry, mesh refinement, CHT, LES, etc.

<u>Technical Accomplishments</u>

- Adding detail to the CFD engine models has provided significant improvement in accuracy over the full operating range without tuning for individual cases
- Mechanism detail, mesh refinement shown to improve accuracy of predicted combustion metrics & emissions
- Full-geometry model better captures charge motion and turbulence providing improved CO accuracy

<u>Collaborations</u>

- Multiple collaborative efforts with OEMs, ISVs, and NLs
- Leverage of DOE's ASCR leadership computing resources at OLCF User Facility
- Future Work (subject to change based on availability of funding and allocation of HPC resources)
 - Complete implementation of CHT model, add LES turbulence model and assess impact over full DoE
 - Use LES to examine impact of flow field perturbations on predicted knock
 - Implement CHT for identification of hotspots and autoignition locations

Dean Edwards: edwardskd@ornl.gov

Charles Finney: finneyc@ornl.gov



Backup Slides



GPUs are enablers for the use of detailed chemistry

LLNL's Zero-RK GPU-enabled chemistry solver provides significant computational speed-up

- For reduced mechanism (144 species)...
 - 33% speed-up of chemistry solution with Zero-RK and GPUs
 - Overall computation time reduced 10-20%
- Simulations with detailed mechanism (766 species) not practical without HPC and GPUs
 - On conventional, CPU-based clusters, after ~15 days only completed from IVC to +10 °ATDC
 - On Titan with Zero-RK and GPUs, ~5 days for IVC to EVO
- Impact of GPUs even greater when increasing cell count (mesh refinement)



Comparison of computational time for chemistry and transport solvers for example case of constant-volume, single-zone model with 1000 cells using CPUs only. GPUs should provide an additional 1.25-2x speed up.



- Emissions index error-squared for HC and NOx across calibration space
 - Note: HC and NOx error plots not on same scale
- Highlights regions of speed-load map where additional improvement needed
 - Caution needed as there are several points with similar speed-load but different engine operating parameters as part of calibration space





Significant improvement observed for most combustion and emissions metrics for individual cases and globally for increased coverage of calibration space without retuning

- Peak cylinder pressure, CA50: Consistent improvement
- HC: Significant improvement with chemistry and wall-spray refinements
- NOx: Significant improvement with chemistry and grid refinements
- CO and soot: Need further work



ASME ICEF2017-3631



Progressive improvement in NOx emissions prediction

- Significant improvement noted with...
 - Chemistry detail: Zel'dovich \rightarrow NOx relations from GRI mechanism
 - Mesh refinement: Better resolution of cell temperature



National Laboratory

RESEARCH CENT

Progressive improvement in emissions prediction over full range

- Overall, absolute soot prediction for many cases became worse
- CONVERGE particulate size mimic (PSM) soot model is linked with SAGE chemistry solvers
 - Not accessible with Zero-RK UDF
 - Runs without Zero-RK also showed increased soot with reduced mechanism





Progressive improvement in emissions prediction over full range

• CO prediction still a significant challenge



