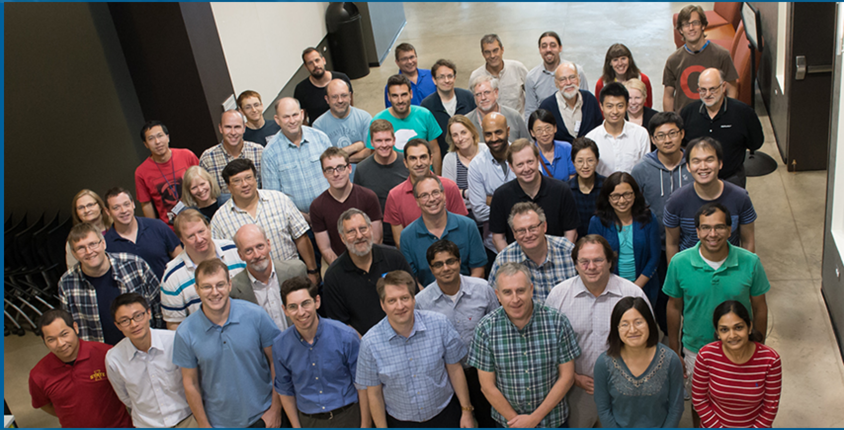
19 JUNE 2017

ISC, FRANKFURT AM MAIN, GERMANY

EXPERIENCES WITH KNL IN THE ALCF EARLY SCIENCE PROGRAM



DAVID E. MARTIN
Manager, Industry Partnerships and Outreach
Argonne Leadership Computing Facility

TIM WILLIAMS
Deputy Director of Science
Argonne Leadership Computing Facility

Argonne NATIONAL LABORATORY

ALCF THETA SYSTEM & EARLY SCIENCE PROGRAM (ESP)

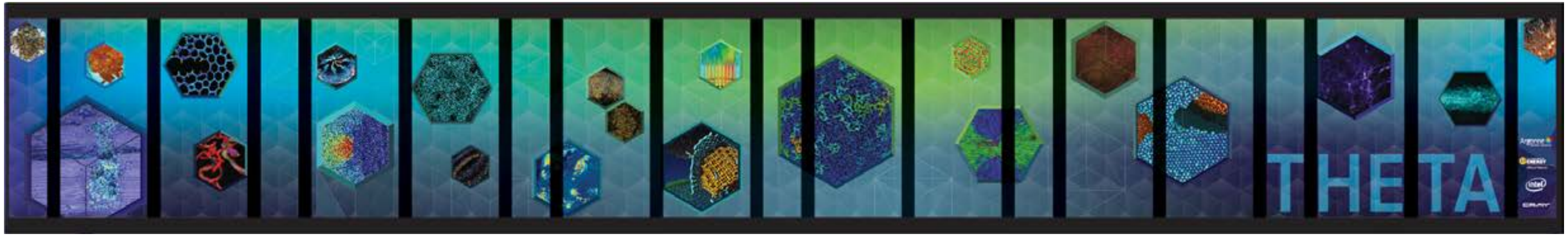
Argonne Leadership Computing Facility

THETA

- 3624 nodes
  - Xeon Phi 7230 (2nd gen.)
  - 16 GB MCDRAM
  - 192 GB DDR4
  - 128 GB SSD
- Peak 9.65 petaFLOPS
- Cray Aries interconnect
- 10 PB Lustre parallel file system

EARLY SCIENCE PROGRAM

- Theta dedicated for science runs: *just ended*
- 6 Tier 1 + 6 Tier 2 projects
- Optimize applications
- Solidify libraries & infrastructure
- Prep Theta for *science on day one*



ALCF Early Science Program

Applications Readiness

- Prepare applications for next-gen system:
  - Architecture
  - Scale
- ~Two year lead time

Proposals

- Ambitious targeted science calculation
- Parallel performance
- Development needed
- Team

Support

PEOPLE

- Funded ALCF postdoc
- Catalyst staff member support
- Vendor experts

TRAINING

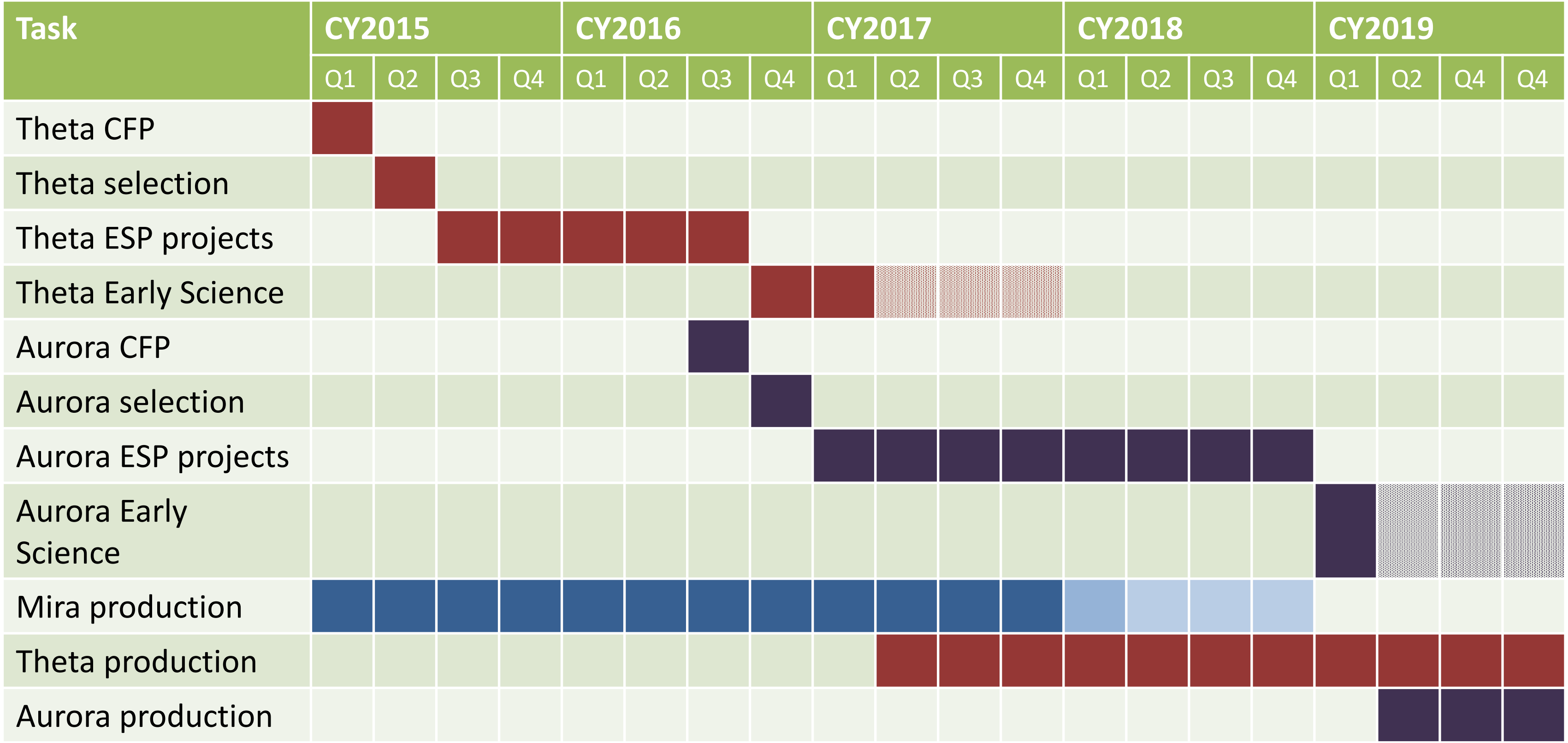
- Training on HW and programming
- Community workshop to share lessons learned

COMPUTE RESOURCES

- Current ALCF systems
- Early next-gen hardware & simulators
- 3 months dedicated Early Science access
  - Pre-production (post-acceptance)
  - Large time allocation
  - Continued access for rest of year

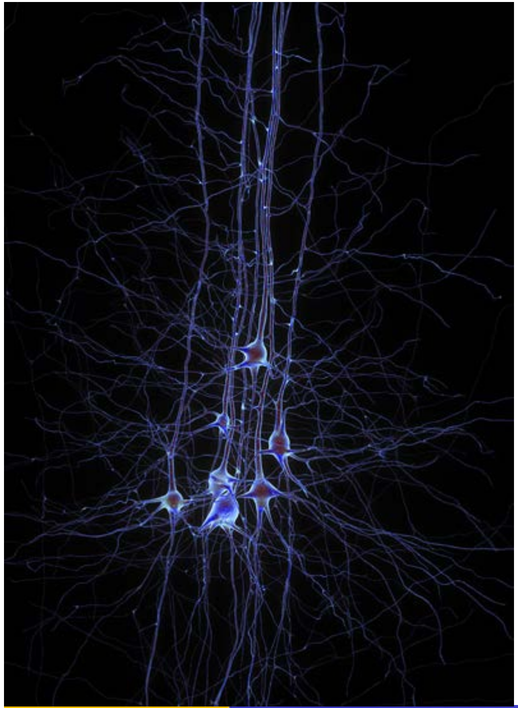
http://esp.alcf.anl.gov

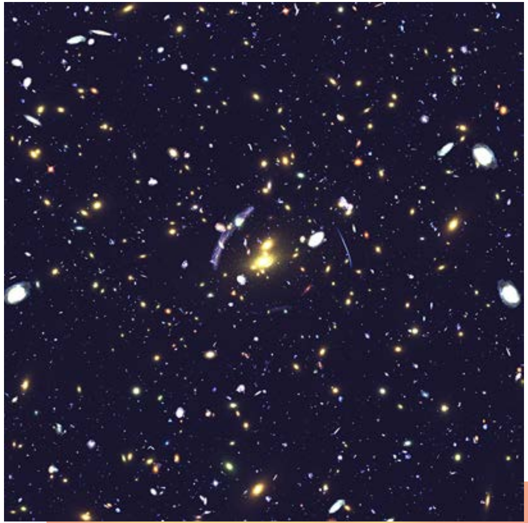
ESP Timeline

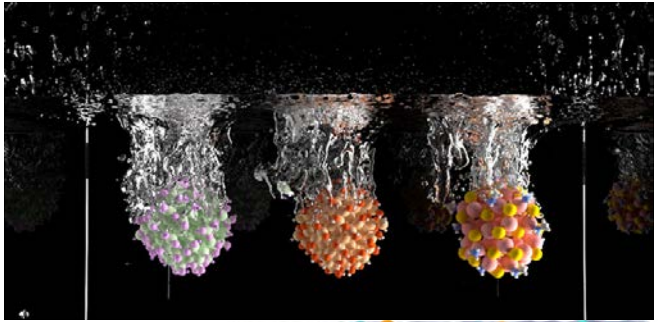


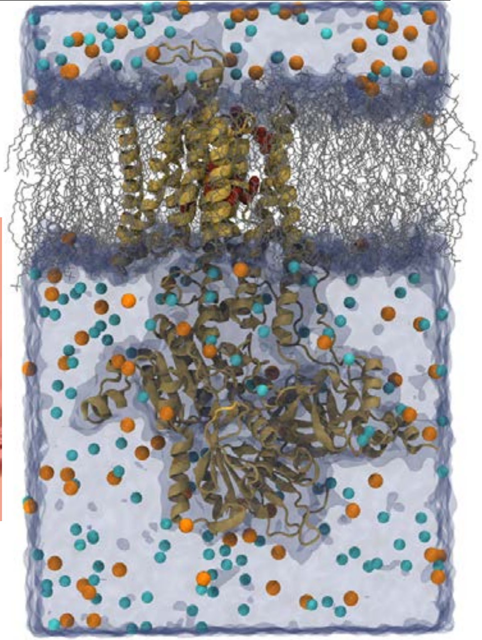
THETA ESP PROJECTS

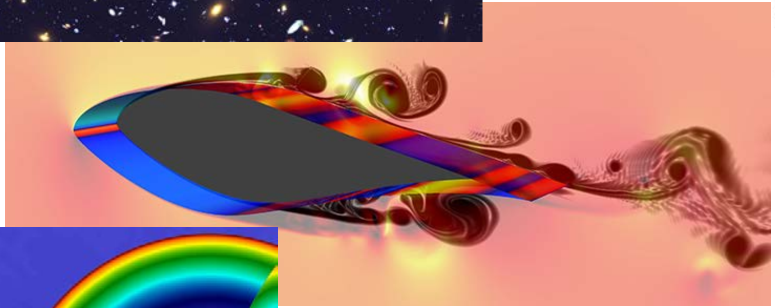
Tier 1

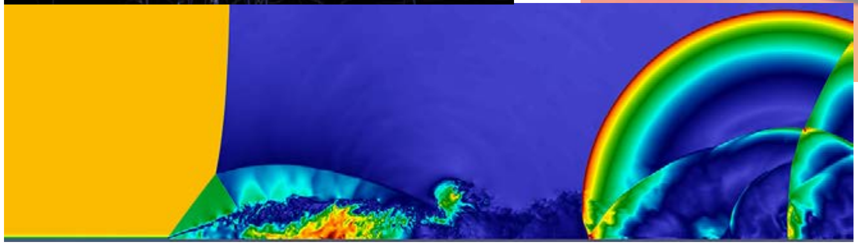












THETA ESP PROJECTS

Tier 1

Code: CoreNeuron
PI: Fabien Delalondre (EPFL)
Many coupled, nonlinear ODEs
Catalysts: Y.Alexeev, T. Williams

Code: HACC
PI: Katrin Heitmann (ANL)
N-body gravity + SPH hydro
Catalysts: H. Finkel, A. Pope
Postdoc: J.D. Emberson

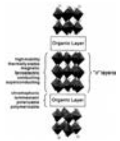
Codes: WEST & Qbox
PI: Giulia Galli (U. Chicago)
MBPT & ab initio MD
Catalyst: C. Knight
Postdoc: H. Zheng

Code: NAMD
PI: Benoit Roux (U. Chicago, ANL)
MD with replica methods
Catalyst: W. Jiang
Postdoc: B. Radak

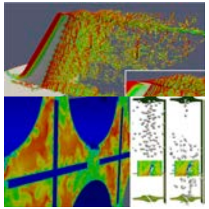
Code: SU2
PI: Juan Alonso (Stanford U)
Large Eddy Simulation, O(3-4)
Catalyst: R. Balakrishnan

Code: HSCD
PI: Alexei Khokhlov (U. Chicago)
DNS, reacting flows, patch AMR
Catalyst: M. Garcia

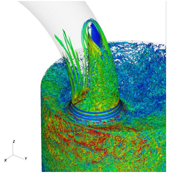
THETA ESP PROJECTS



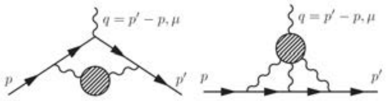
Codes: FHI-Aims & GAtor
PI: Volker Blum (Duke U.)
MBPT (DFT) & genetic algorithm
Catalyst: Álvaro Vázquez-Mayagoitia



Code: PHASTA
PI: Kenneth Jansen (U. Colorado)
CFD, unstructured mesh
Catalyst: Hal Finkel

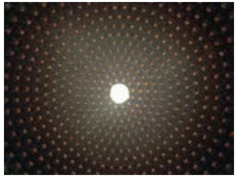


Code: Nek5000
PI: Christos Frouzakis (ETHZ)
Spectral element CFD with combustion
Catalyst: Scott Parker



Codes: MILC & CPS
PI: Paul Mackenzie (FNAL)
Lattice QCD
Catalyst: James Osborn

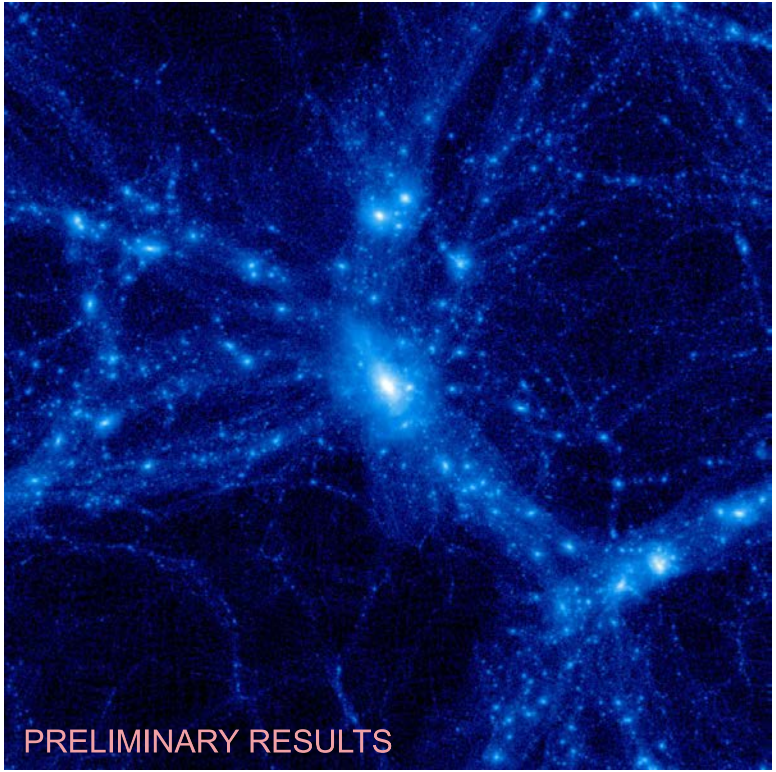
Code: GAMESS
PI: Mark Gordon (Iowa State U.)
FMO - quantum chemistry
Catalysts: Yuri Alexeev, Graham Fletcher



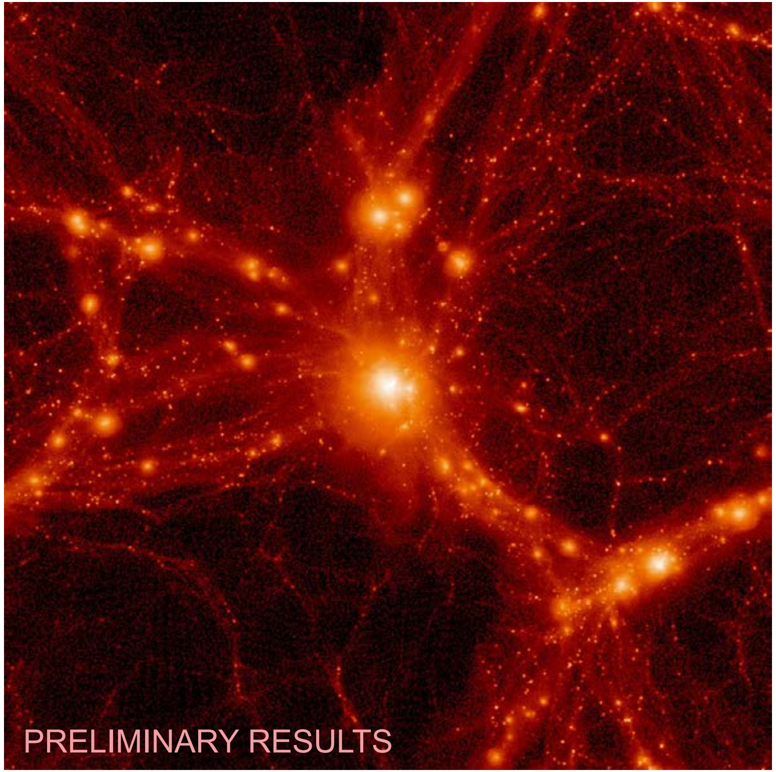
Code: GFMC
PI: Steven Pieper (ANL)
Greens Function Monte Carlo – nuclear
Catalyst: James Osborn

NEXT-GENERATION COSMOLOGY SIMULATIONS WITH HACC: CHALLENGES FROM BARYONS

Code: HACC • *PI: Katrin Heitmann (ANL)* • N-body gravity + CRKSPH hydrodynamics • *Catalysts: Hal Finkel, Adrian Pope* • *Postdoc: J.D. Emberson*



Cold dark matter



Baryons

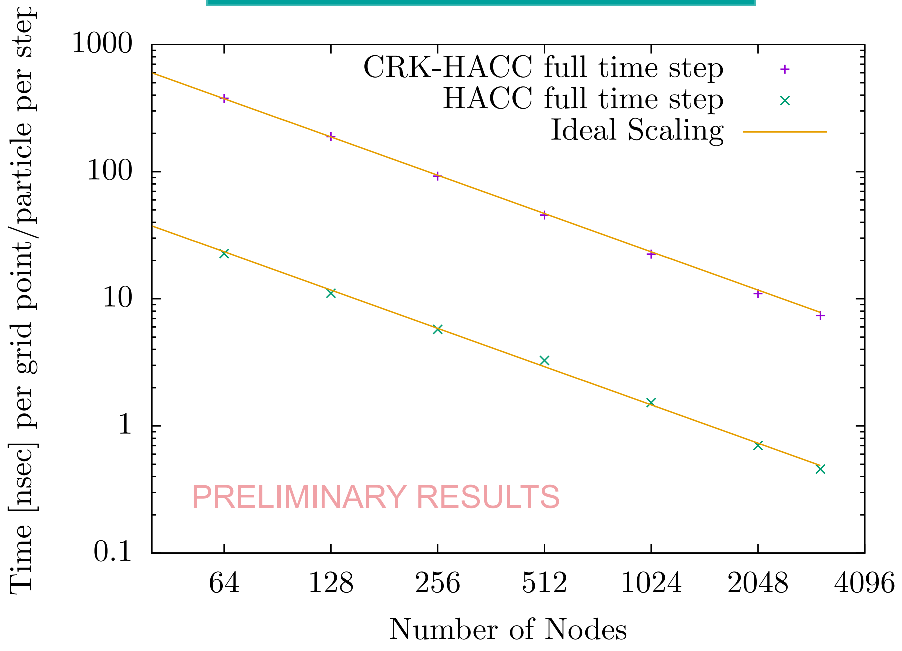
NEXT-GENERATION COSMOLOGY SIMULATIONS WITH HACC: CHALLENGES FROM BARYONS

Code: HACC • *PI: Katrin Heitmann (ANL)* • N-body gravity + CRKSPH hydrodynamics • *Catalysts: Hal Finkel, Adrian Pope* • *Postdoc: J.D. Emberson*

- KNL-specific short-range PP force code
  - Work by Vitali Morozov
  - KNL 6X faster than BG/Q per core
    - 4X wider vectors
    - 0.8X clock rate
    - 2X FP instructions per cycle
- Machine-independent long-range force code

•2048^3 particles on 64 to 2048 nodes (2304^3 particles on 3072 nodes)
•8 ranks per node, 16 threads
•Cache-quad mode

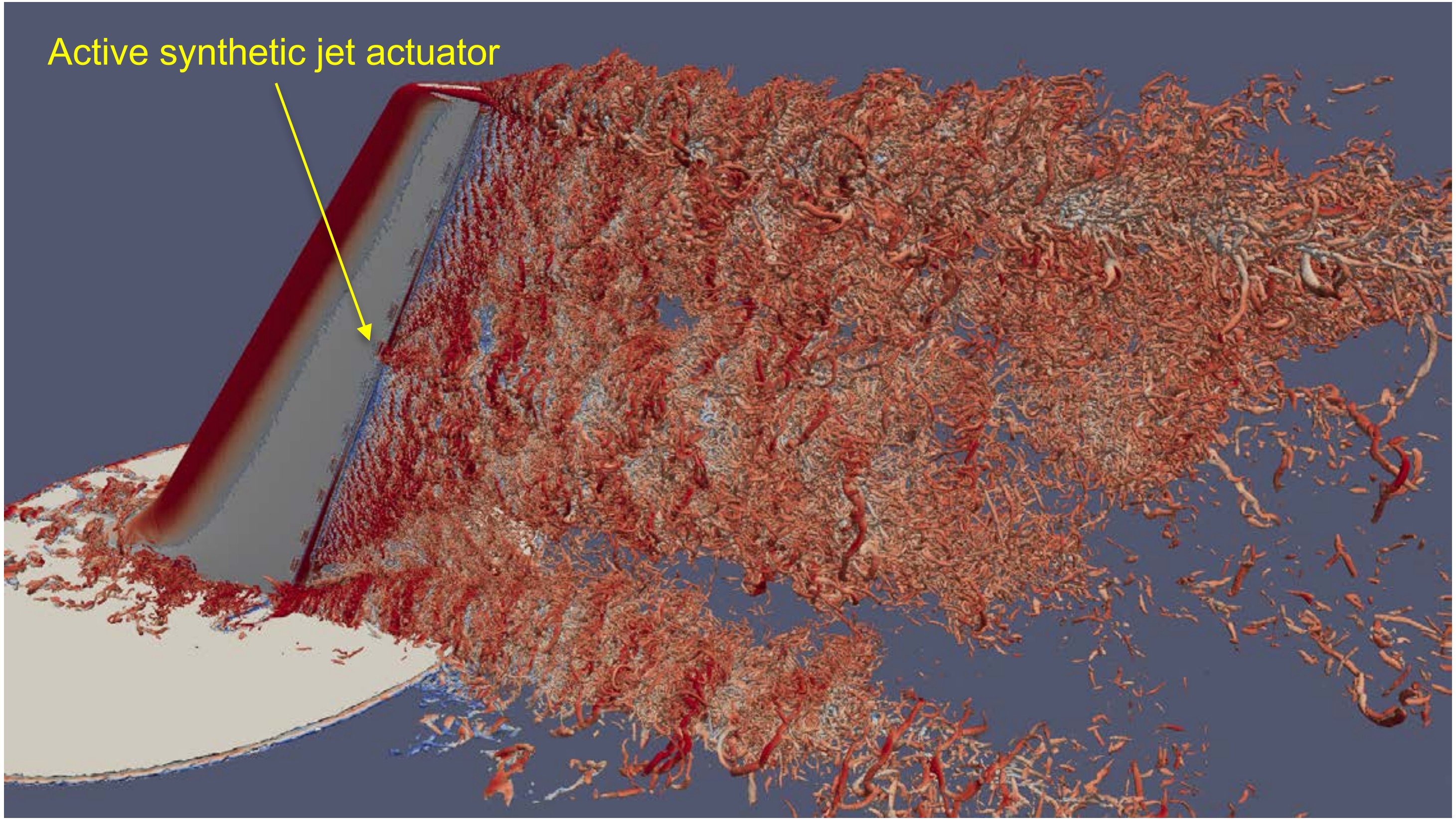
Strong scaling on Theta



Argonne NATIONAL LABORATORY

EXTREME SCALE UNSTRUCTURED ADAPTIVE CFD: AERODYNAMIC FLOW CONTROL

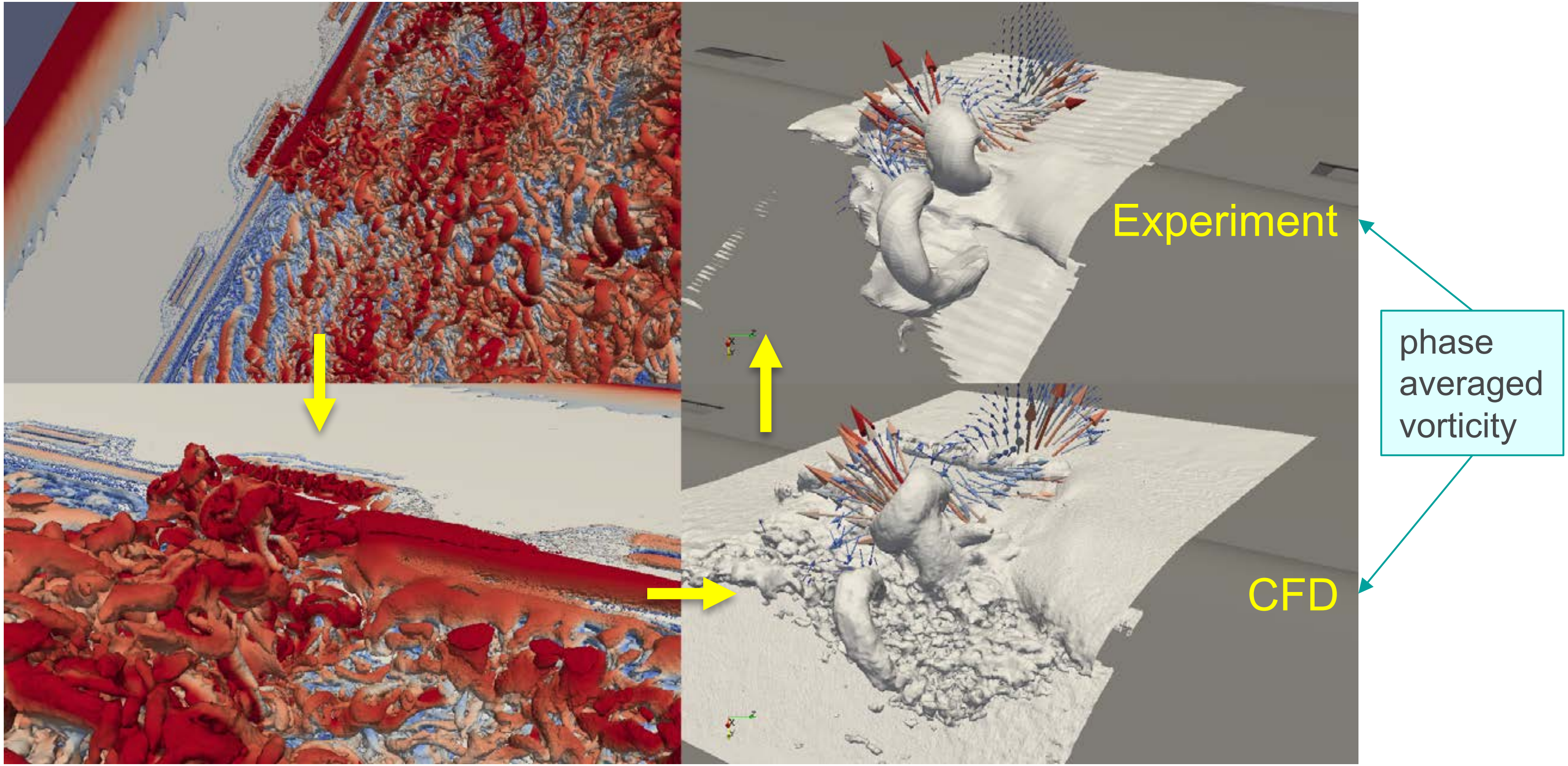
Code: PHASTA • *PI: Kenneth Jansen (U. Colorado Boulder)* • CFD, unstructured mesh • *Catalyst: Hal Finkel*



- 3D finite element
- unstructured adaptive mesh
- fully implicit
- 5 billion elements
- 2048 Theta nodes (128K KNL cores)

EXTREME SCALE UNSTRUCTURED ADAPTIVE CFD: AERODYNAMIC FLOW CONTROL

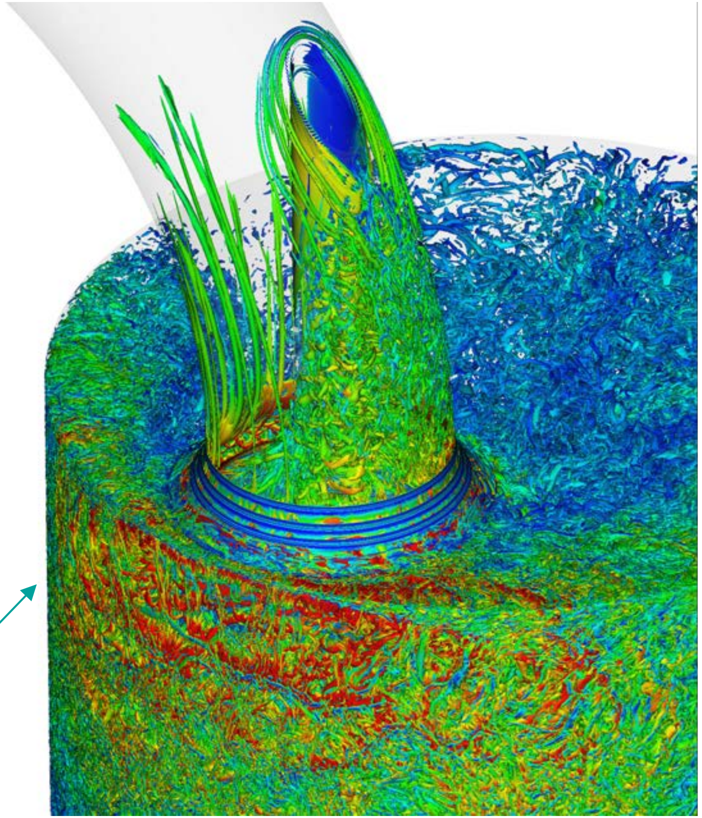
Code: PHASTA • *PI: Kenneth Jansen (U. Colorado Boulder)* • CFD, unstructured mesh • *Catalyst: Hal Finkel*



FLOW, MIXING AND COMBUSTION OF TRANSIENT TURBULENT GASEOUS JETS IN CONFINED CYLINDRICAL GEOMETRIES

Code: Nek5000 • *PI: Christos Frouzakis (ETHZ)* • Spectral element CFD with combustion • *Catalyst: Scott Parker*

- Internal combustion engine
  - Compression ignition
  - Natural gas as fuel
- Direct numerical simulation
- Low Mach number combustion
- Strong scale to fit in MCDRAM

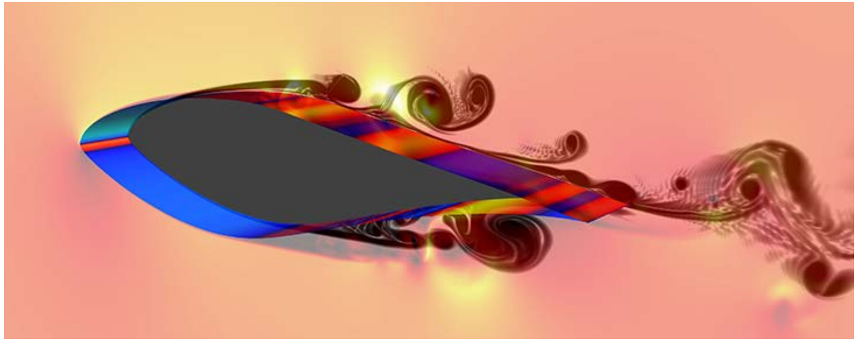


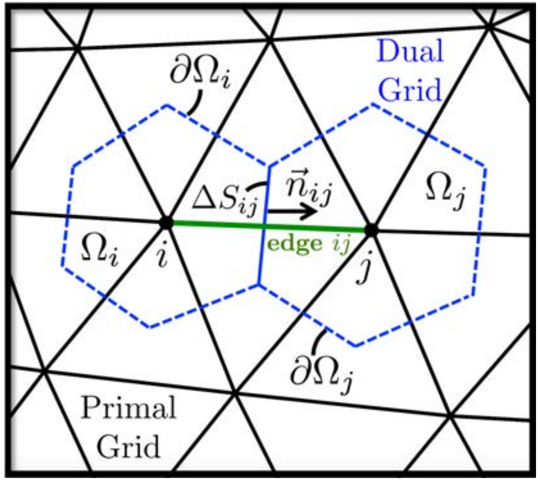
Turbulence in intake stroke of ICE simulated on Theta

SCALE-RESOLVING SIMULATIONS OF WIND TURBINES WITH SU2

Code: SU2 • *PI: Juan Alonso (Stanford U.)* • CFD (LES) • *Catalyst: Ramesh Balakrishnan*

- Large Eddy Simulation (LES) of a few turbines plus tower
  - Third order finite volume
  - High order discontinuous Galerkin
- LES results feed reduced-order Kinematic Simulation for wind farm design
- SU2 evolving into high-end open source CFD package (community code)
  - Finite volume methods
  - Unstructured mesh

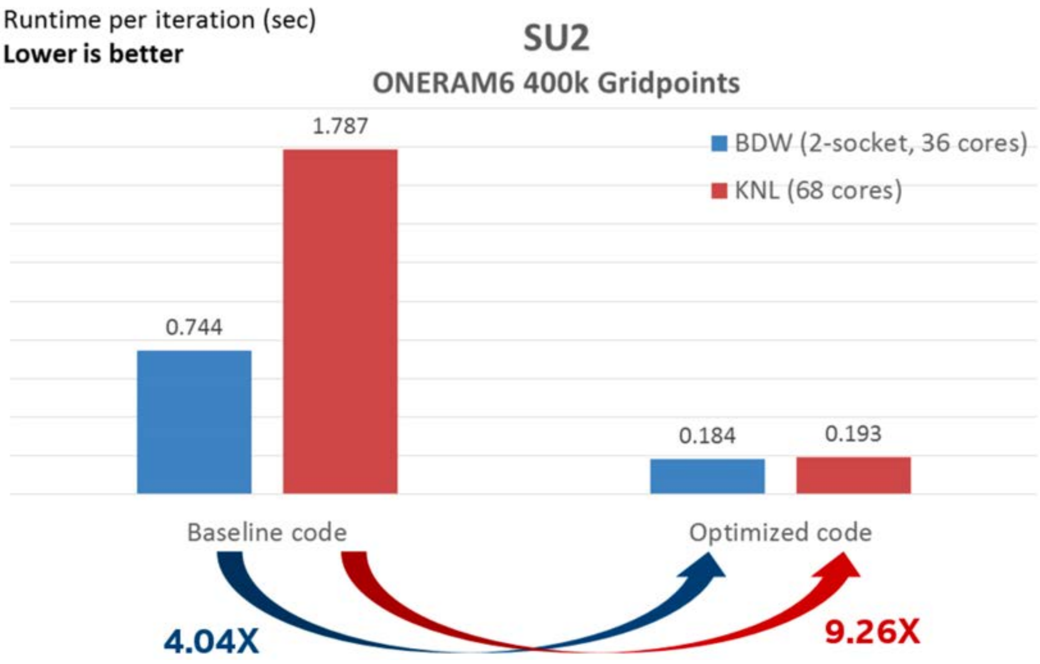




Gaurav Bansal, Anand Deshpande, Dheevatsa Mudigereb, others (Intel); Tom Economon, Francisco Palacios (Stanford)

SCALE-RESOLVING SIMULATIONS OF WIND TURBINES WITH SU2

Optimizations



- Threads
  - Single OMP parallel region at high level in program
- Vectorization
  - Outer loop (edges)
  - Loop tiling
  - Reduced gathers-scatters
- Memory hierarchy
  - Edge/vertex RCM reordering
  - Smart allocation
  - Change AoS to SoA

Gaurav Bansal, Anand Deshpande, Dheevatsa Mudigereb, others (Intel); Tom Economon, Francisco Palacios (Stanford)

THETA ESP LESSONS LEARNED

- Structure of Arrays
- Strong scale to fit in MCDRAM
  - Successes with many MPI ranks per node (up to 64)
- Transition from BGQ (MPI + OpenMP) → KNL not generally painful
  - Adjust ranks/threads sweet spot
- Memory access looks like streaming?
  - `#pragma vector nontemporal`
- Use MKL FFT (multiple electronic structure codes)

THETA ESP (AND OTHER) LESSONS LEARNED

- Running within MCDRAM? Cache mode as good as flat mode
  - Flat mode: `numactl -m 1` (allocate in HBM; error if spills out)

Code	Method	Runtime flat vs. cache
HACC	Tree N-body, particle-mesh	-0.1%
WEST	Many body perturbation theory	+8.98%
Qbox	Ab initio molecular dynamics	-6.6%
USQCD	Several Lattice QCD methods	Virtually no difference
NAMD	Classical molecular dynamics	No significant difference
QMCPACK	Quantum Monte Carlo	-4.8%
VSVB	Electronic structure	+4.2%, +0.59%