On the Path to Exascale: Reactor Neutronics in ECP

Tom Evans
HPC Methods and Applications Team
ECP Applications Development, Energy Portfolio

July 27, 2020

Outline

- ECP Overview
  - DOE HPC Roadmap
- ECP ExaSMR Project
  - SMR Challenge Problem
  - Code products
  - Monte Carlo neutronics challenges
  - Intra-node performance
  - On-the-fly Doppler broadening
  - Inter-node domain decomposition strategies
  - Beyond ECP

ECP Overview

ECP by the Numbers

7
YEARS

$1.8B

A seven-year, $1.8B R&D effort that launched in 2016

6
CORE DOE
LABS

Six core DOE National Laboratories: Argonne, Lawrence Berkeley, Lawrence Livermore, Oak Ridge, Sandia, Los Alamos

- Staff from most of the 17 DOE national laboratories take part in the project

3
FOCUS
AREAS

Three technical focus areas: Hardware and Integration, Software Technology, Application Development supported by a Project Management Office

80+
R&D TEAMS

1000
RESEARCHERS

More than 80 top-notch R&D teams

Hundreds of consequential milestones delivered on schedule and within budget since project inception

The three technical areas in ECP have the necessary components to meet national goals

Performant mission and science applications @ scale

Aggressive RD&D Project	Mission apps & integrated S/W stack	Deployment to DOE HPC Facilities	Hardware tech advances

Application Development (AD)

Develop and enhance the predictive capability of applications critical to the DOE

24 applications including national security, to energy, earth systems, economic security, materials, and data

Software Technology (ST)

Deliver expanded and vertically integrated software stack to achieve full potential of exascale computing

70 unique software products spanning programming models and run times, math libraries, data and visualization

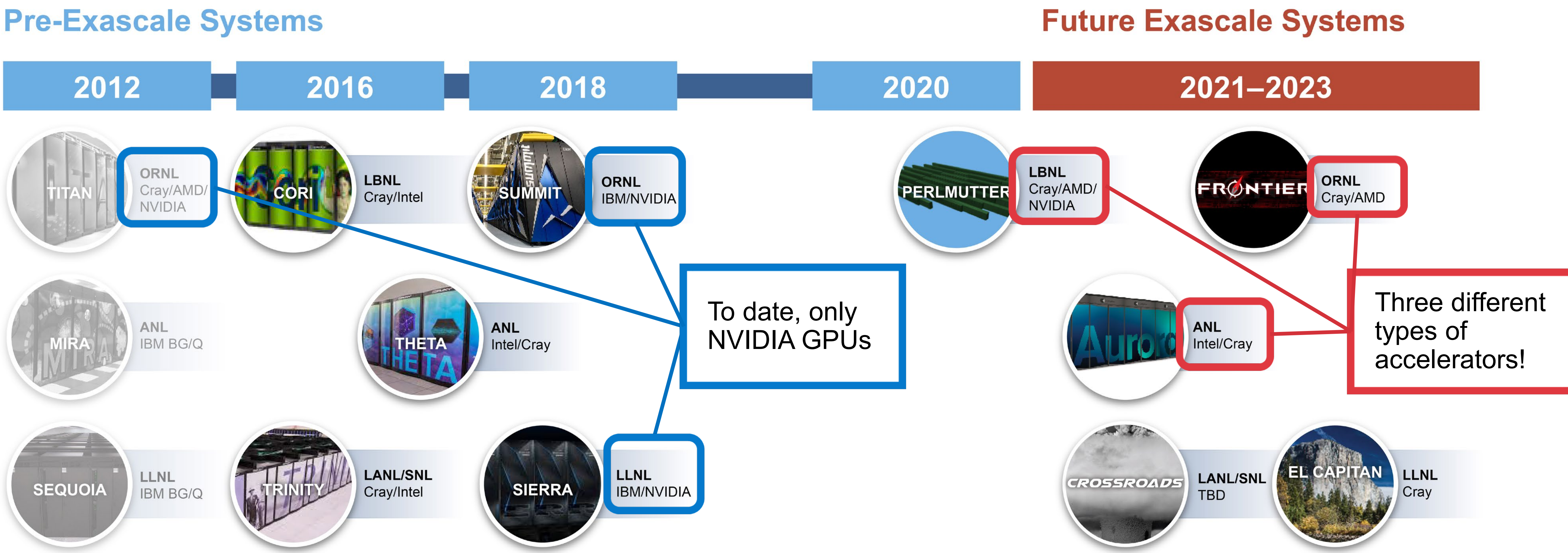
Hardware and Integration (HI)

Integrated delivery of ECP products on targeted systems at leading DOE HPC facilities

6 US HPC vendors focused on exascale node and system design; application integration and software deployment to facilities

Department of Energy (DOE) Roadmap to Exascale Systems

An impressive, productive lineup of *accelerated node* systems supporting DOE's mission



ExaSMR: Coupled Monte Carlo Neutronics and Fluid Flow Simulation of Small Modular Reactors

EXASCALE COMPUTING PROJECT

Partner institutions

Project Budget: $21,370,811
Actual Cost: $10,221,425

Institution	Site PI	Investigation Areas
Oak Ridge National Laboratory	Steven Hamilton	Neutronics, Multiphysics
Argonne National Laboratory	Paul Romano	Neutronics, Multiphysics, CFD
MIT	Kord Smith	Reactor physics
Penn State University	Elia Merzari	CFD

Oak Ridge National Laboratory

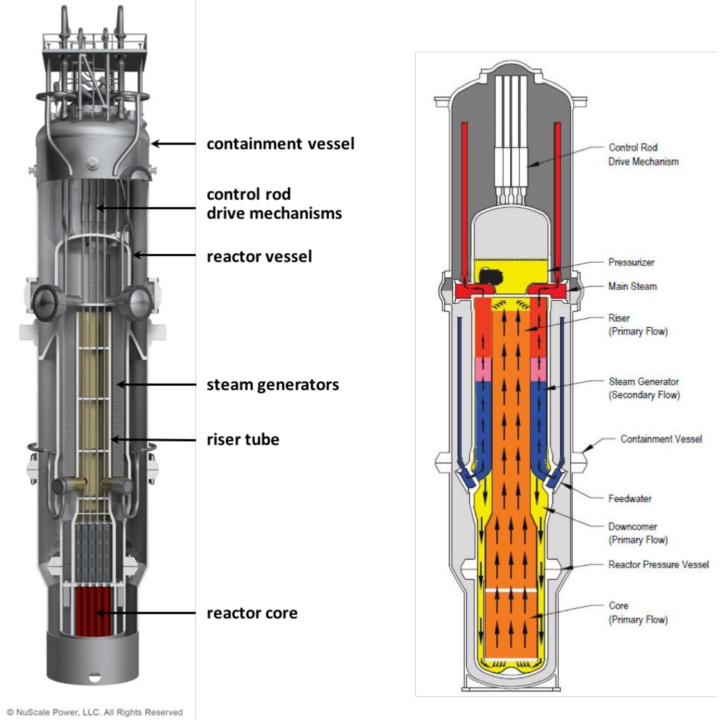
Argonne National Laboratory

Massachusetts Institute of Technology

PennState

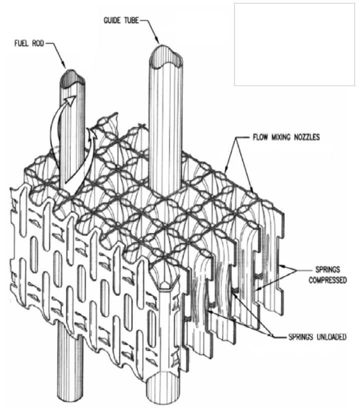
ExaSMR: Modeling and Simulation of Small Modular Reactors

- Small modular nuclear reactors present significant simulation challenges
  - Small size invalidates existing low-order models
  - Natural circulation flow requires high-fidelity fluid flow simulation
- ExaSMR will couple most accurate available methods to perform “virtual experiment” simulations
  - Monte Carlo neutronics
  - CFD with turbulence models



Reproduced with permission from NuScale Power

MC Neutronics		CFD	
Petascale	**Exascale**	**Petascale**	**Exascale**
• System-integrated responses • Single physics • Constant temperature • Isotopic depletion on assemblies • Reactor startup	• Pin-resolved (and sub-pin) responses • Coupled with T/H • Variable temperatures • Isotopic depletion on full core • Full-cycle modeling	• Single fuel assembly • RANS • Within-core flow	• Full reactor core • Hybrid LES/RANS • Entire coolant loop

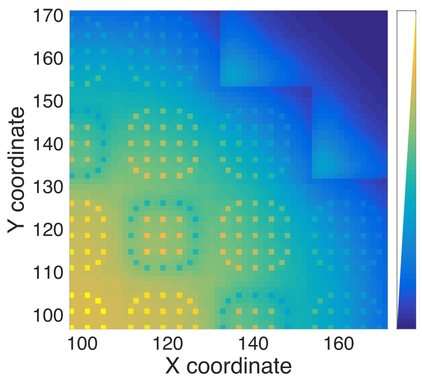


Fuel assembly mixing vane

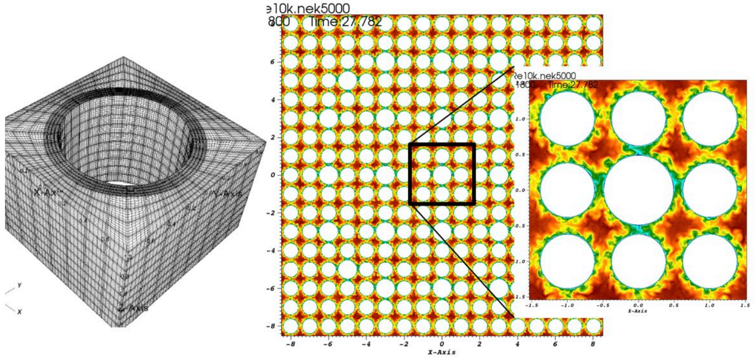
ExaSMR Exascale Challenge Problem

Challenge Problem

- Simulation of full NuScale SMR model core by coupling continuous-energy Monte Carlo neutronics with CFD
  - Complete in-vessel coolant loop
  - Hybrid LES/RANS turbulence model
  - Sub-pin resolution fission power
  - Isotopic depletion (quasi-static)



Total reaction rate in SMR core

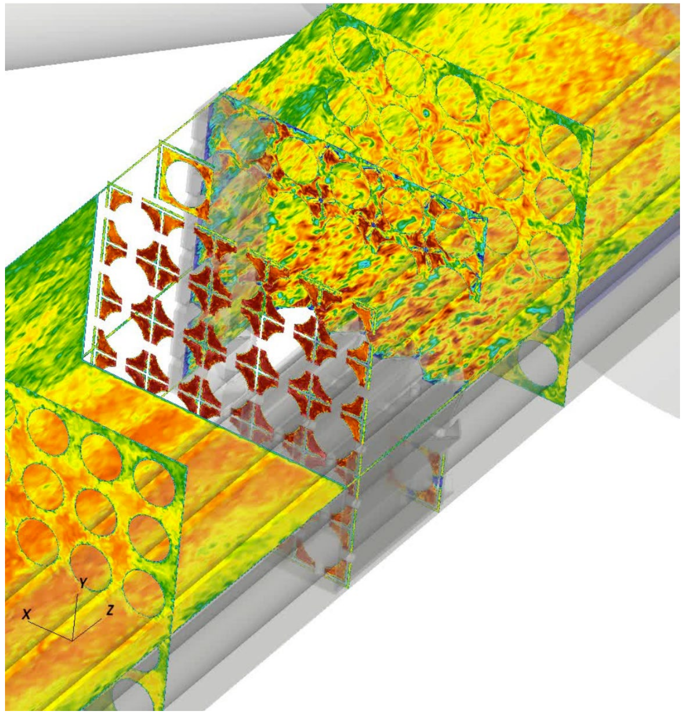

Velocity profile for flow within nuclear fuel assembly

Problem Parameters

- Neutronics
  - Full core representative SMR model containing 37 assemblies with 17×17 pins per assembly and 264 fuels pins per assembly
  - 10^{10} particles per eigenvalue iteration
  - Pin-resolved reaction rate with 3 radial tally regions and 20 axial levels
  - O(150) nuclides and O(8) reactions per nuclide in each tally region
- CFD
  - Assembly bundle mesh models with momentum sources from a resolved CFD calculation on a representative spacer grid
  - Full core mesh 40×10^6 elements and 22×10^9 degrees-of-freedom

NekRS – Nek5000 for advanced computing architectures

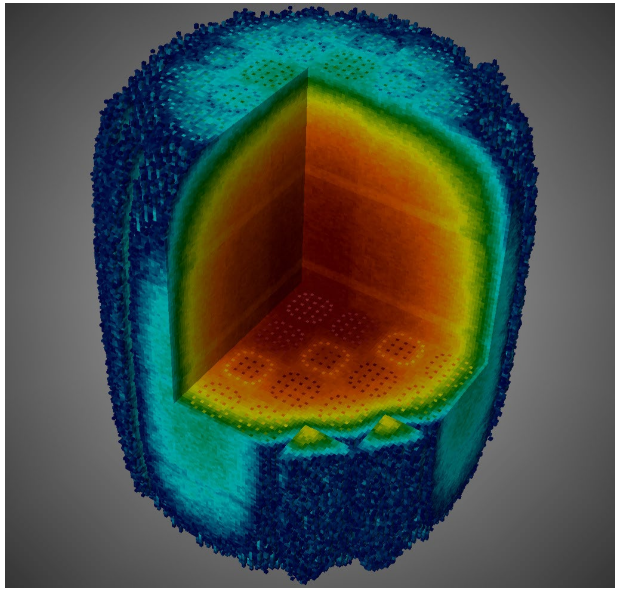
- Nek5000 is spectral finite element CFD solver
  - RANS and LES turbulence models
  - 1999 Gordon Bell Prize winner
- NekRS is revamped version of Nek5000
  - Achieved 4x performance improvement over “native” Nek5000 port using OpenACC
  - Developed in collaboration with ECP CEED co-design project
  - Supports diverse set of computing architectures via libParanumal library
- Full core reactor models currently in development/testing



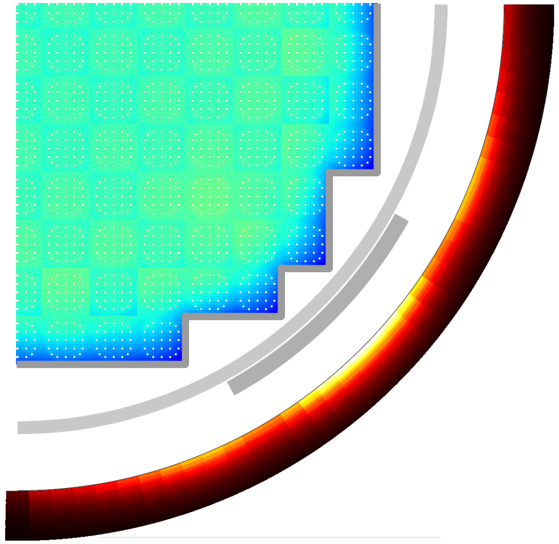
Coolant flow through mixing vane

Shift – MC transport for NVIDIA and AMD GPUs

- Continuous-energy MC radiation transport code in ORNL SCALE suite
  - Heavily used in DOE CASL project for pressure vessel fluence and dosimetry
  - Supports reactor physics and radiation shielding workloads



Total neutron interaction rate in SMR core

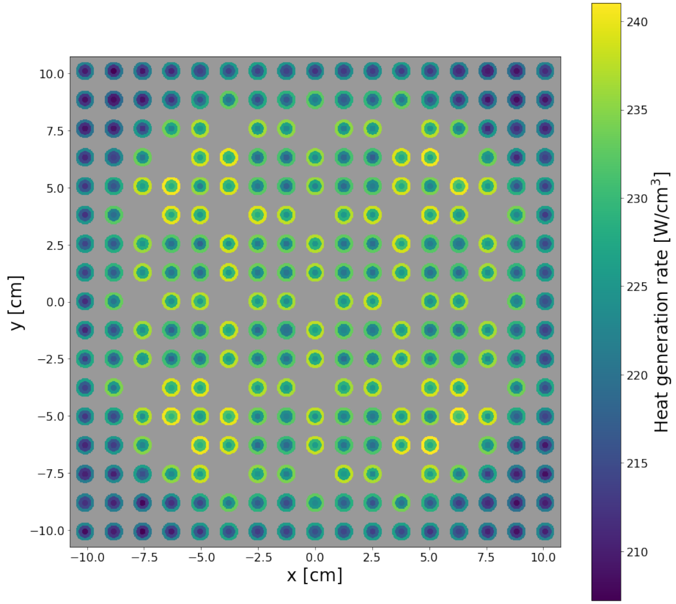


Shift vessel fluence calculation

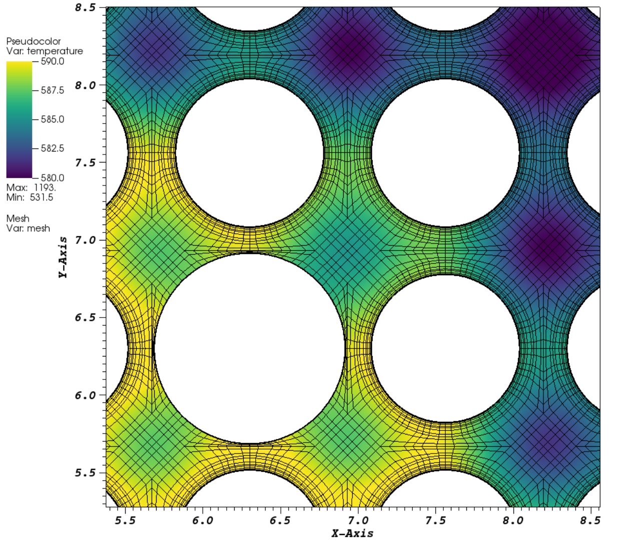
- First production MC transport code to run on GPUs
  - Ported to NVIDIA GPUs using CUDA to utilize Summit
  - Support for AMD GPUs planned by converting CUDA kernels to HIP
  - Allows continued use of NVIDIA GPUs on Summit while developing for Frontier

ENRICO – Simplified multiphysics driver

- ExaSMR has developed ENRICO as a targeted tool for multiphysics coupling
  - Manages program flow and parallel data transfer
  - Supports multiple physics implementations, including low-order models for testing
- Implements communication patterns scalable to very large node counts
- Scaled to full 3D assembly simulation
  - Excellent agreement between different code implementations



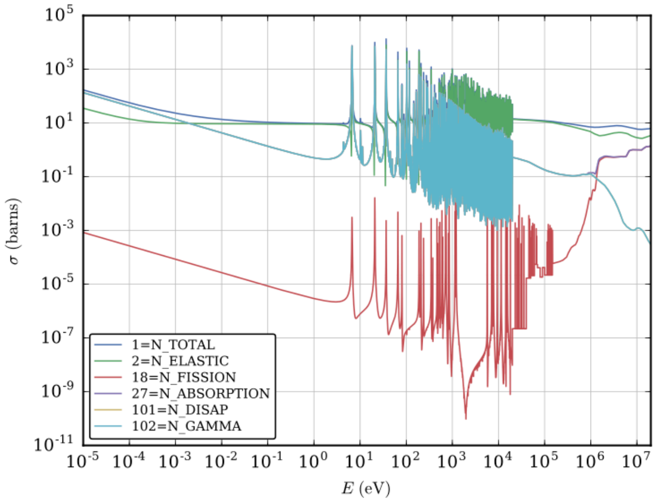
Heat generation rate within fuel assembly

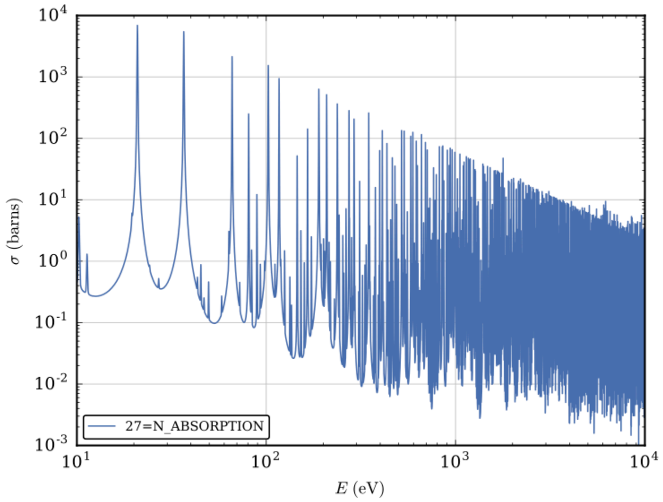


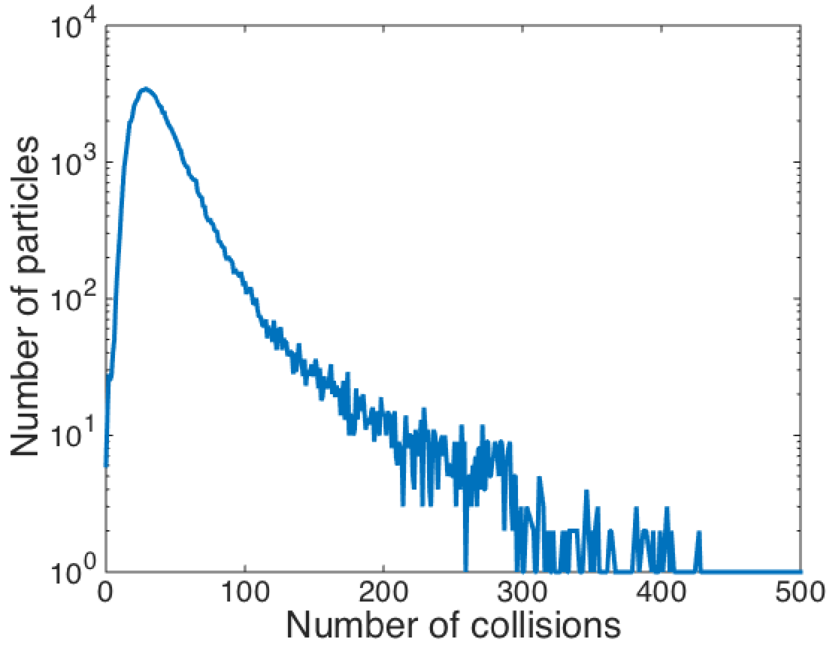
Detail of coolant temperature around fuel pins

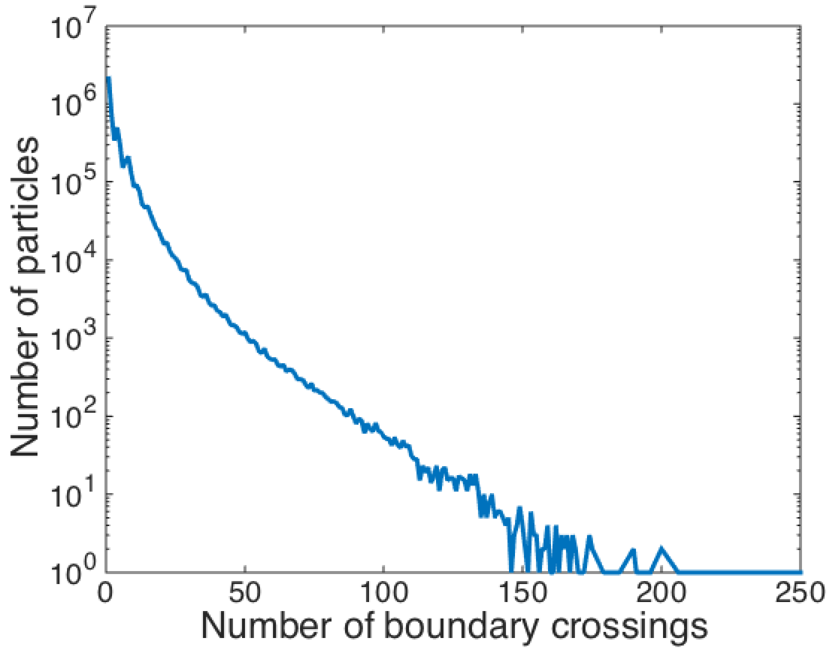
Monte Carlo Neutron Transport Challenges

- MC neutronics is a stochastic method
  - Independent random walks are not readily amenable to SIMT algorithms – on-node concurrency
  - Sampling data is randomly accessed
  - Sampling data is characterized by detailed structure
  - Large variability in transport distributions both within and between particle histories





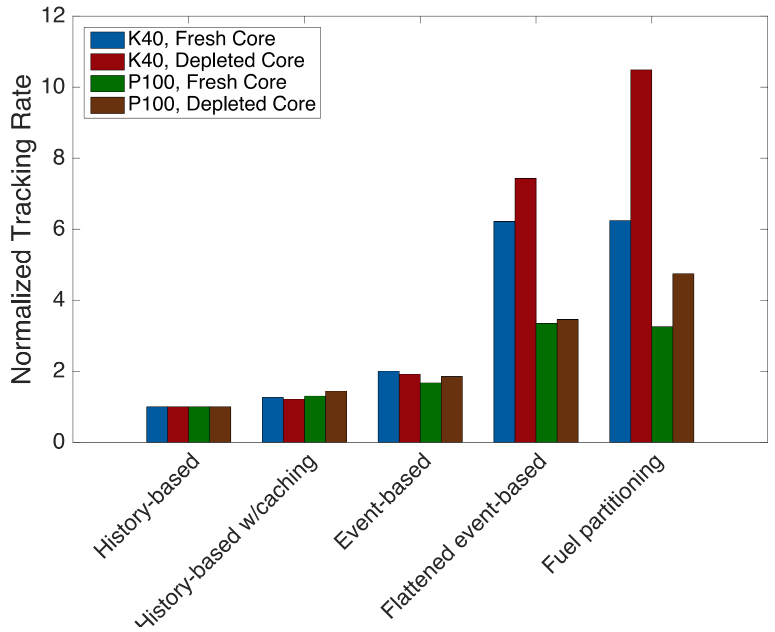


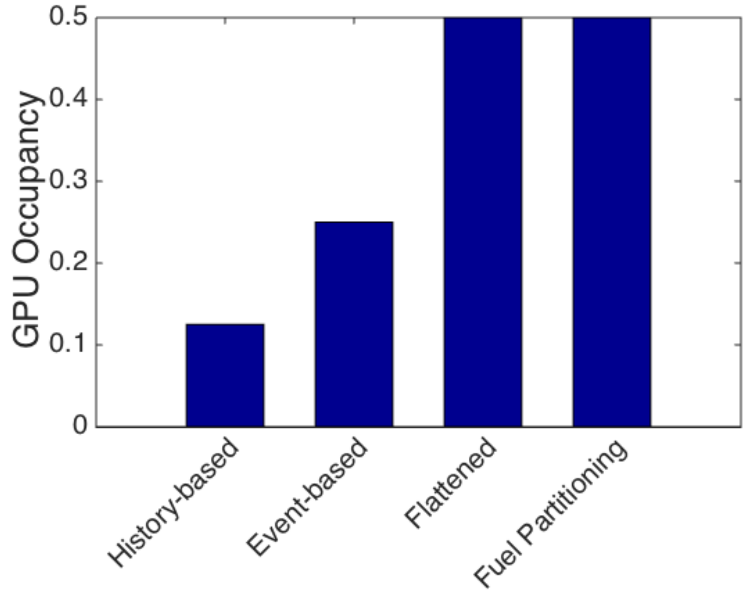


EXASCALE COMPUTING PROJECT

Developing GPU Continuous Energy Monte Carlo – Intra-Node

- Focus on high-level thread divergence
- Optimize for device **occupancy**
  - Separate geometry and physics kernels to increase occupancy
  - Boundary crossings (geometry)
  - Collision (physics)
- Smaller kernels help address variability in particle transport distributions - latency
- Partition macro cross section calculations between fuel and non-fuel regions – separate kernels for each
- Use of hardware atomics for tallies and direct sort addressing
- Judicious use of *texture* memory
  - __ldg on data interpolation bounds

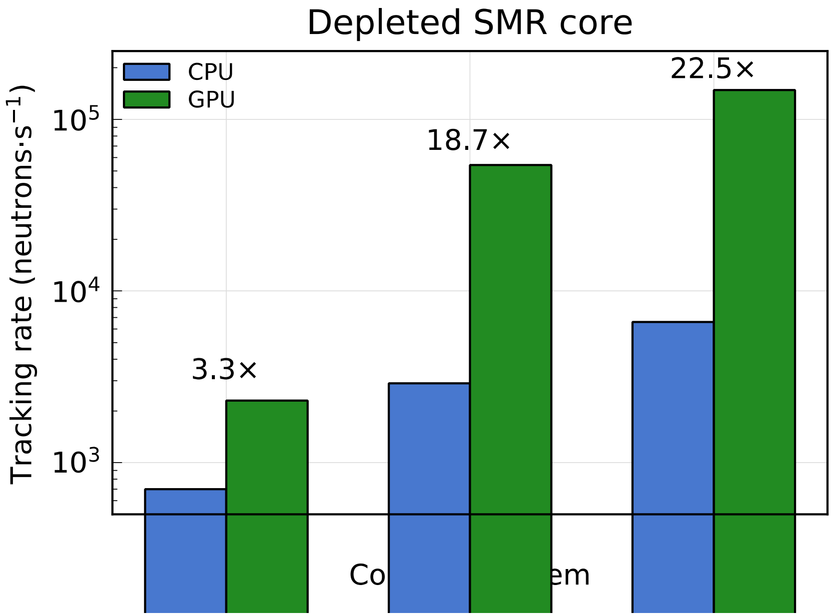


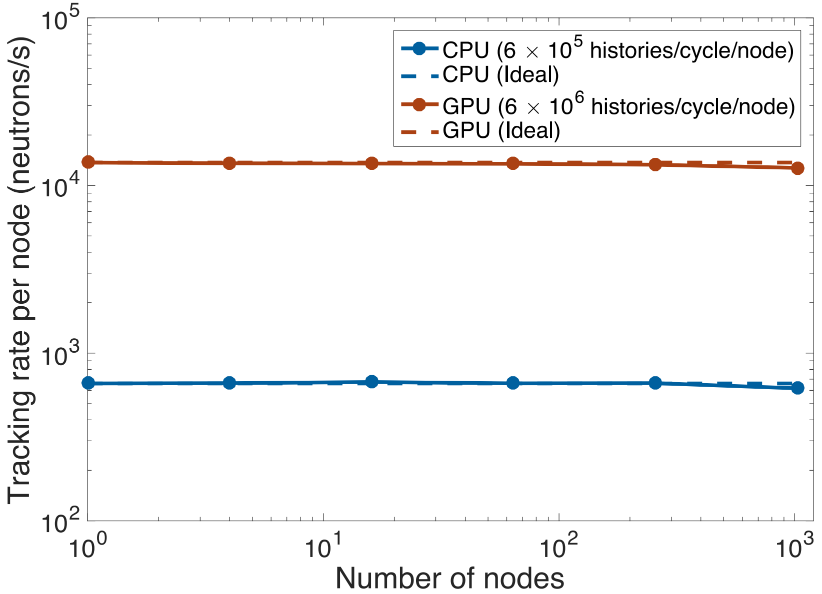


EXASCALE COMPUTING PROJECT

Summit performance results

- Summit system at ORNL has provided a valuable progress assessment
- CFD solver achieved a 7x speedup over Titan using only 15% of Summit
  - Expected 48x improvement at full machine
- MC solver achieved 23x performance increase on nearly 90% of Summit
- Both codes are outpacing increases in machine theoretical peak
  - Algorithmic improvements enable more efficient use of new machines

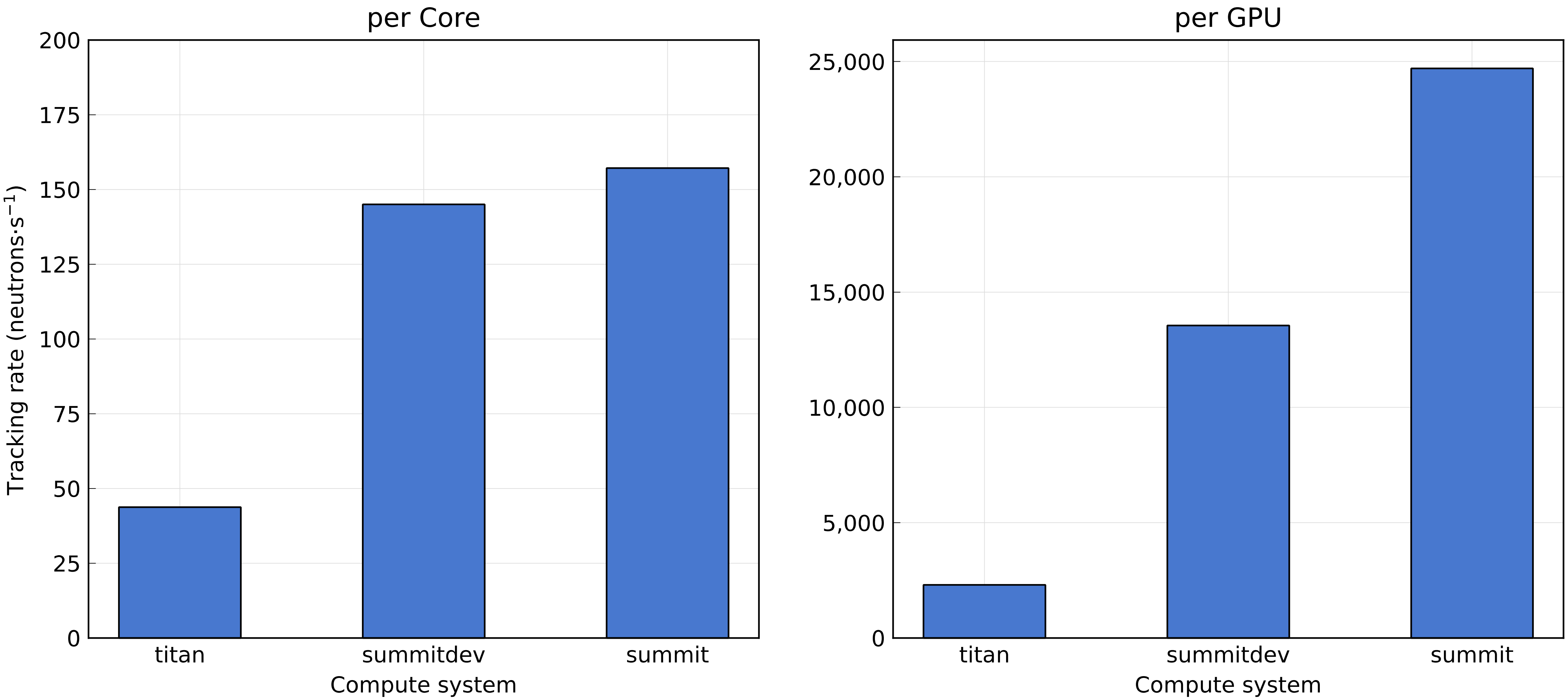




Hamilton, S.P., Evans, T.M., 2019. Continuous-energy Monte Carlo neutron transport on GPUs in the Shift code. *Annals of Nuclear Energy* **128**, 236 – 247. https://doi.org/10.1016/j.anucene.2019.01.012

EXASCALE COMPUTING PROJECT

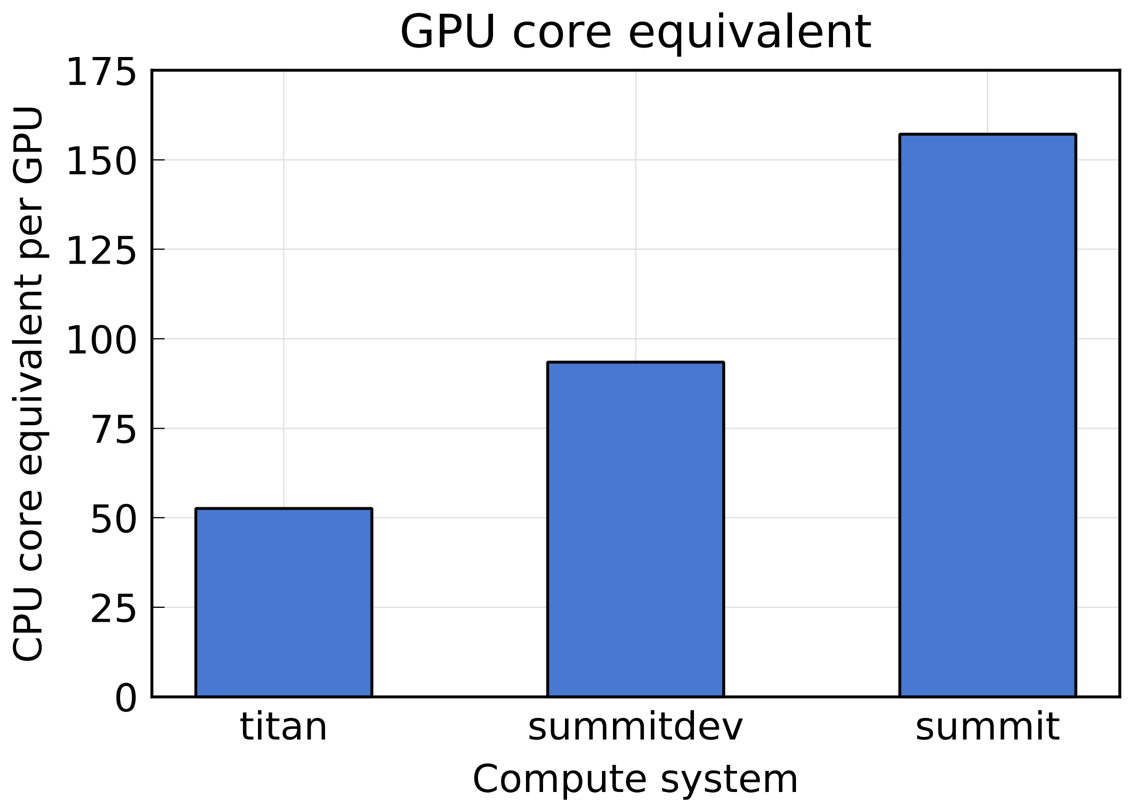
GPU versus CPU generational improvements

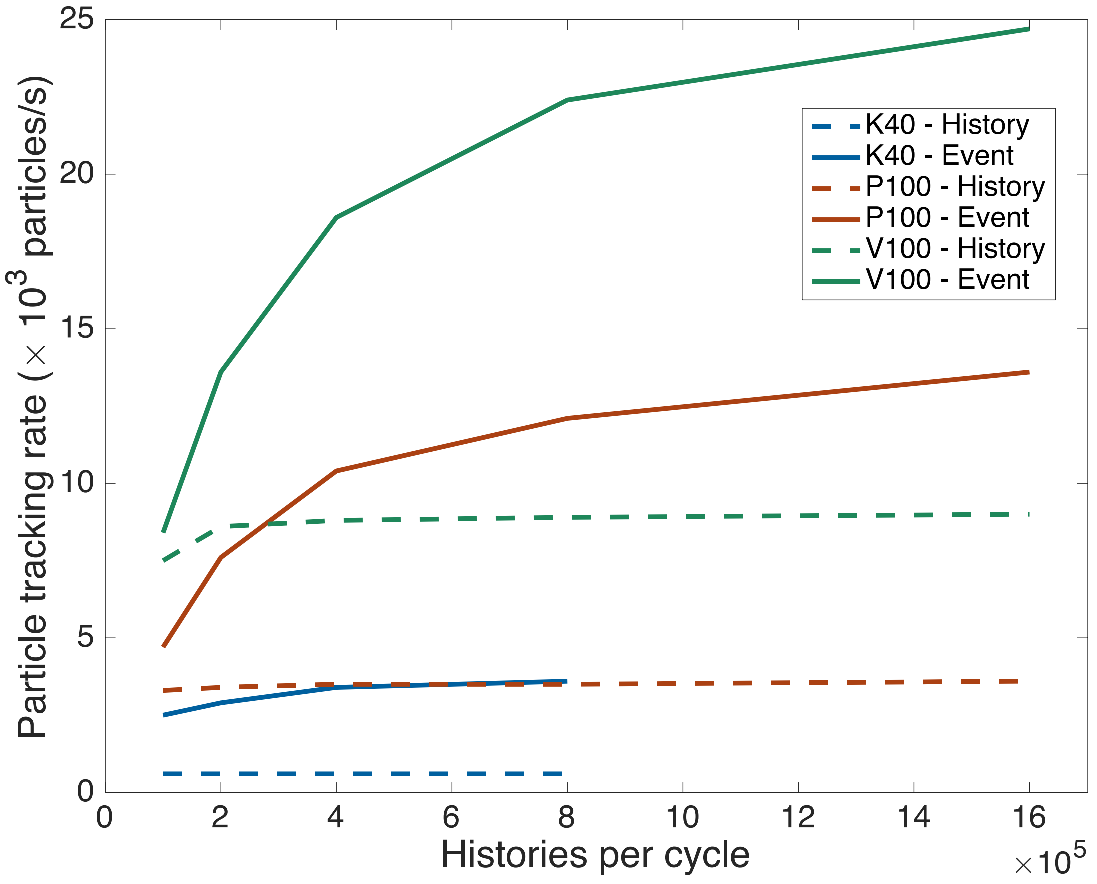


EXASCALE COMPUTING PROJECT

Device saturation and performance

Newest architectures remain unsaturated at 1M particles per GPU

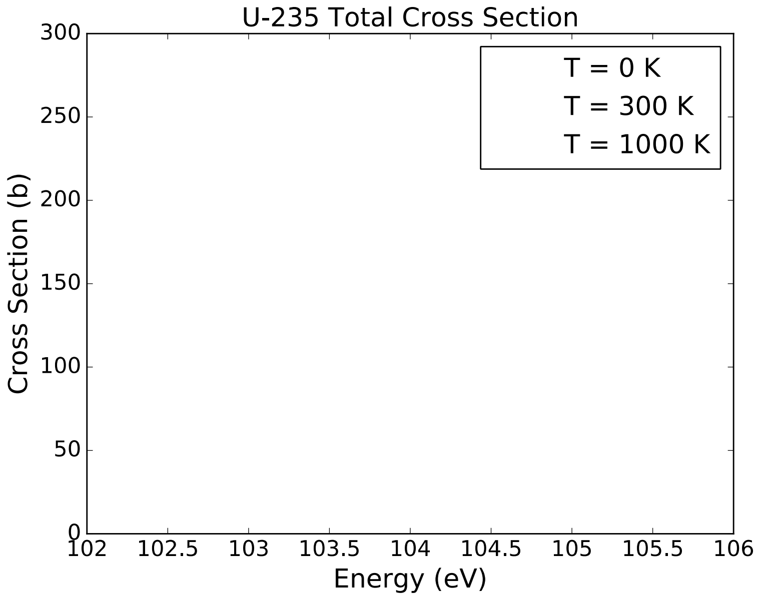




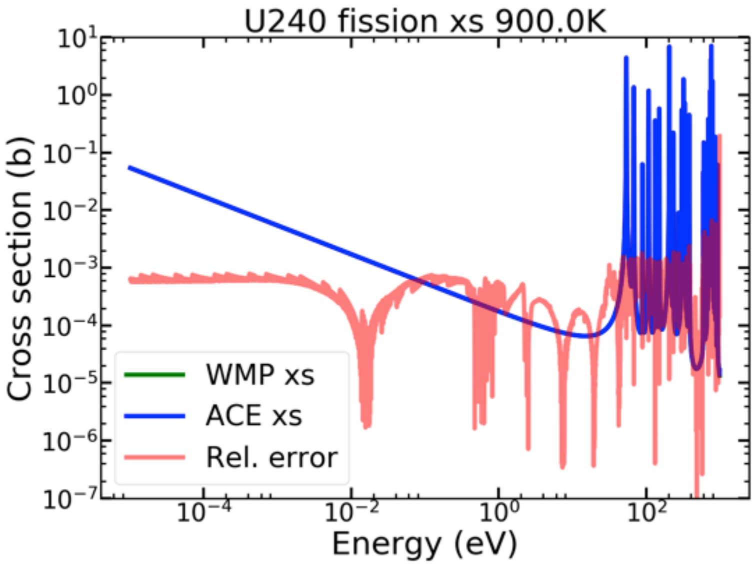
Depleted SMR core

On-the-Fly Doppler Broadening

- Cross section data is sensitive to temperature changes
  - Resonances are flattened as temperature increases
  - Significant impact on reactor operation
- No consensus in broader community about correct approach to treat Doppler broadening “on the fly” in MC transport
  - Needed for multiphysics simulations
- ExaSMR took the windowed multipole method from theory all the way to a production implementation
  - Data processing implemented in ORNL’s AMPX nuclear data utility



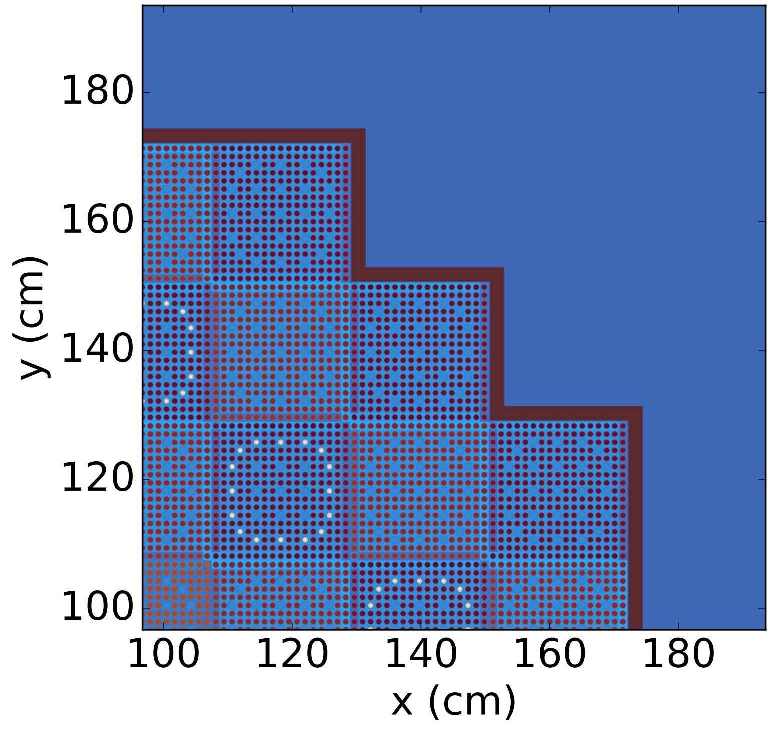
Broadening of U-235 resonances

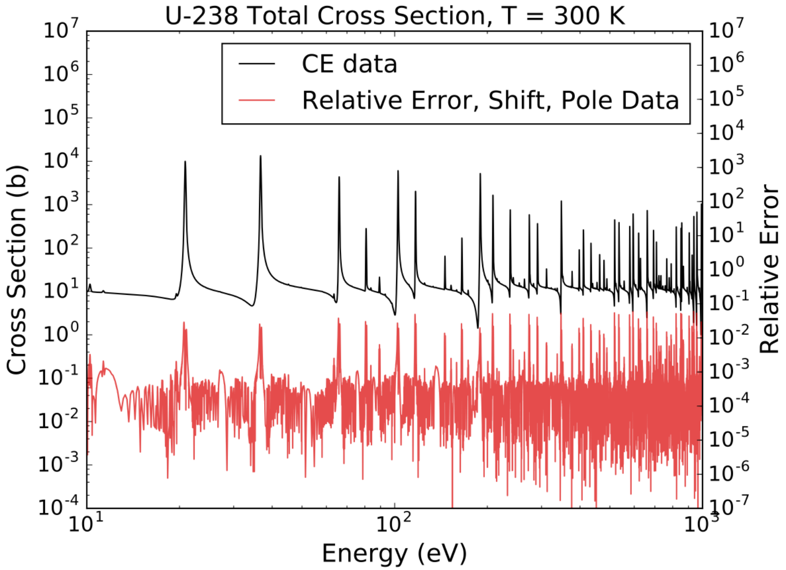


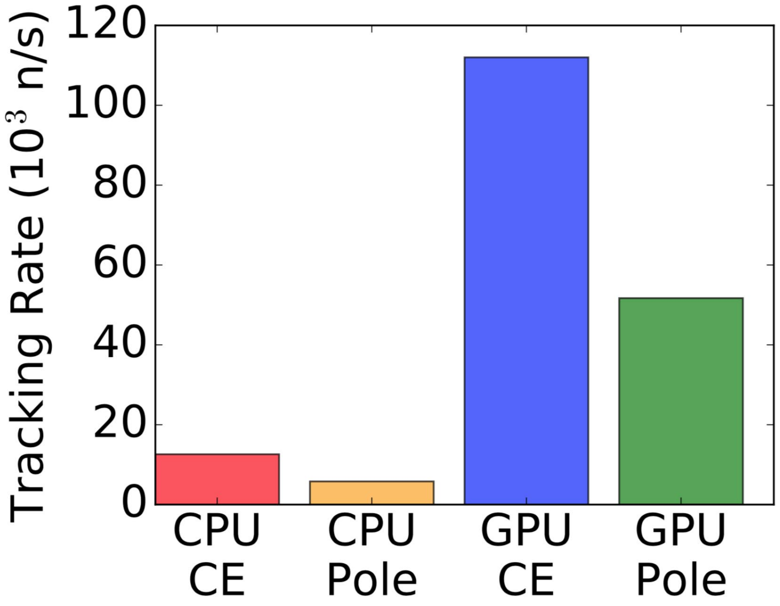
WMP fitting of U-240 fission data

GPU performance

- Performance testing with a quarter-core model of the awaited NuScale Small Modular Reactor (SMR)
- No significant sacrifice of accuracy compared to standard continuous energy (CE) data
- Each GPU thread does individual Fadeeva evaluations (no vectorization over nuclides)
- Factor of 2-3 performance penalty on both the CPU and GPU using Pole Method for Doppler Broadening



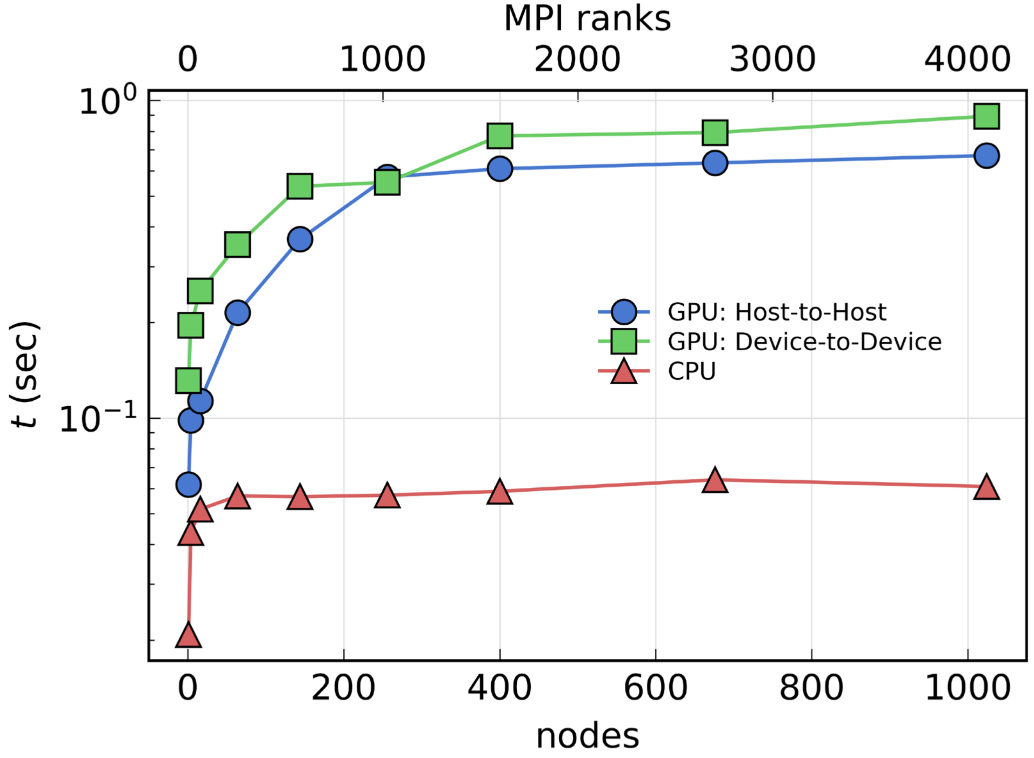




2x IBM Power8+
4x NVIDIA Tesla P100

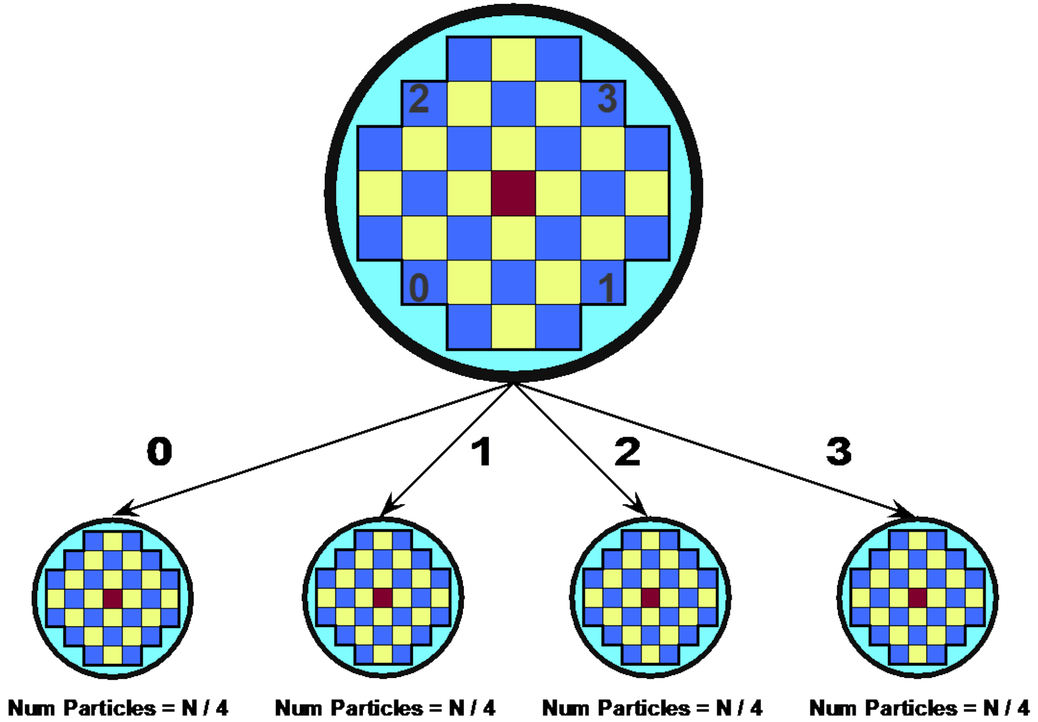
Inter-node Scaling

2 Resource Sets per Node (2 GPUs per Resource Set)

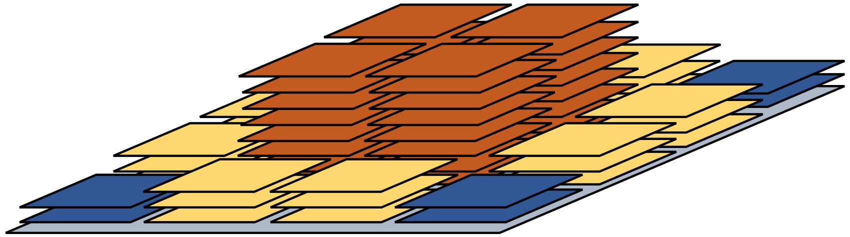


Investigating MPI-aware CUDA

- Communication device-to-device (bypass NIC)
- Does not currently give same performance as manually moving data
- Next-gen platforms will optimize device-to-device



Multi-set domain decomposition topology
(in development – GPU)



Intra-set non-uniform block out to address load balancing

EXASCALE COMPUTING PROJECT

Ellis, J.A., Evans, T.M., Hamilton, S.P., Kelley, C.T., Pandya, T.M., 2019. Optimization of processor allocation for domain decomposed Monte Carlo calculations. *Parallel Computing* **87**, 77–86. https://doi.org/10.1016/j.parco.2019.06.001

Collaboration efforts beyond ECP

- Algorithmic improvements in Nek5000 and Shift directly benefit DOE NEAMS
- Naval Nuclear Laboratory personnel attend ExaSMR and ECP project meetings
- Ongoing collaboration with NNSA labs (*Summit on Summit* working group)
- Team members contributed to OLCF Summit machine acceptance and early science campaign
- ExaSMR tools are being leveraged in a GAIN voucher proposal with X-energy
- ExaSMR Monte Carlo technologies are being leveraged by HEP (ORNL/FNAL) for advance particle physics transport on GPUs

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

NEAMS NUCLEAR ENERGY ADVANCED MODELING & SIMULATION PROGRAM

CASL

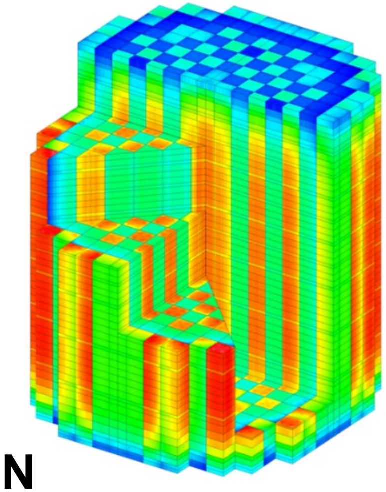
NAVAL NUCLEAR LABORATORY

Fermilab

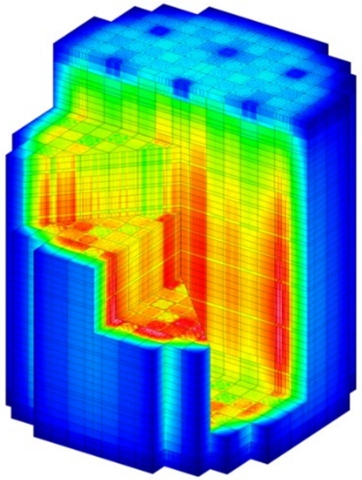
NNSA National Nuclear Security Administration

ECP EXASCALE COMPUTING PROJECT

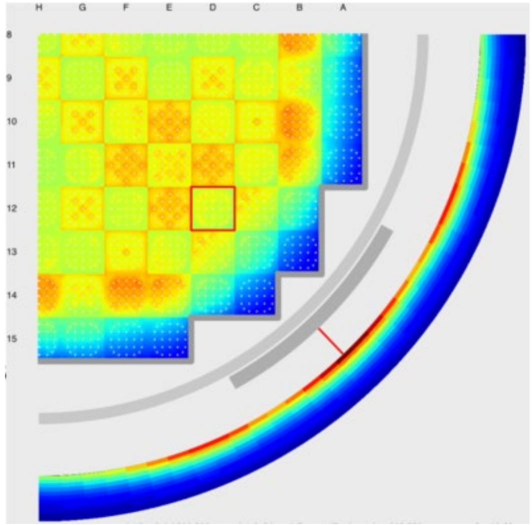
ECP Connections to NE Programs – NEAMS Applications



ORIGEN
Isotopic depletion and decay

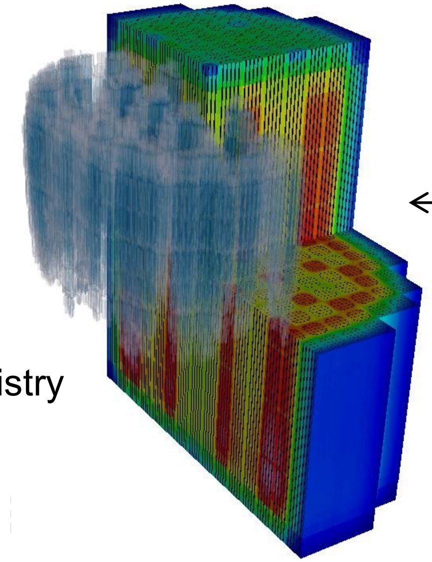


MPACT
Reactor physics

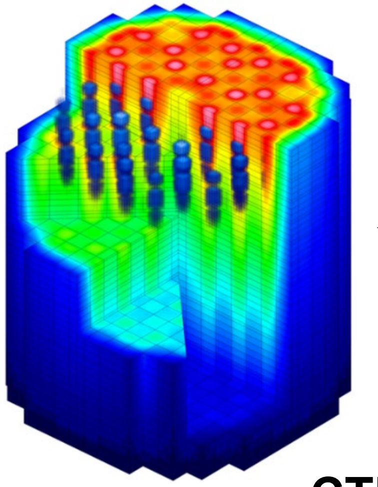


Shift
Monte Carlo Neutron and Gamma transport

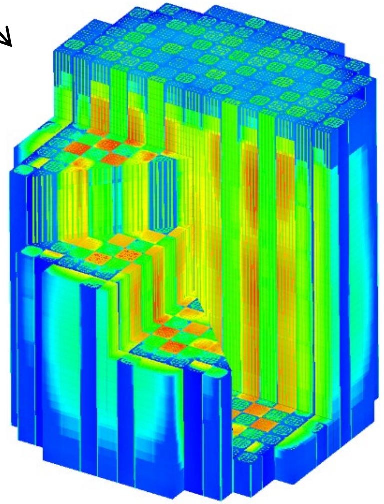
Shift is primary vehicle for excore mod-sim in NEAMS program



MAMBA
CRUD chemistry



CTF
Thermal-hydraulics

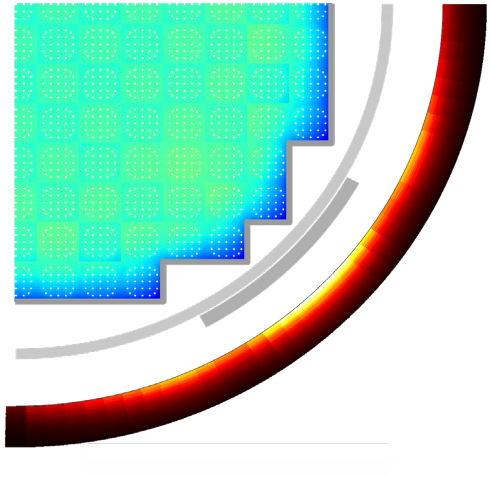


BISON
Fuel rod performance

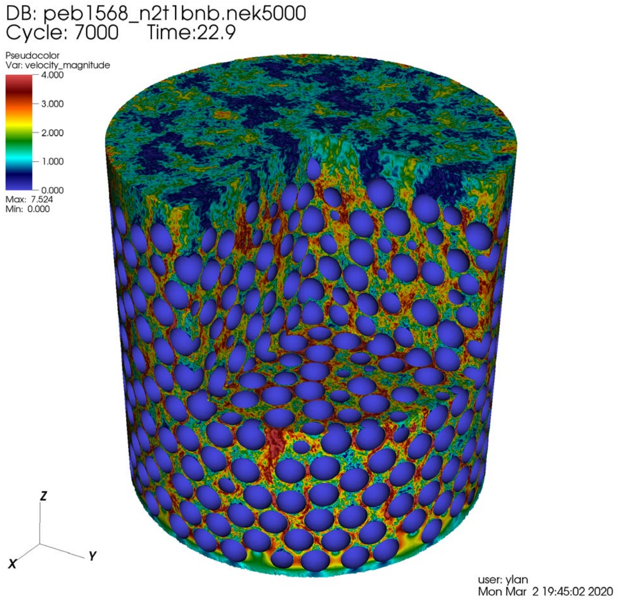
VERA Physics Codes

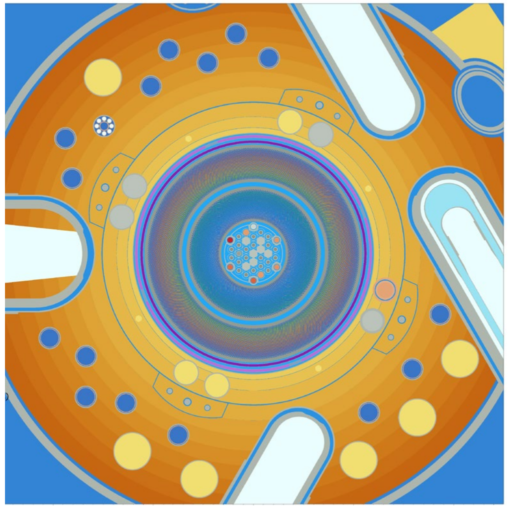
Courtesy: Dave Kropaczek, CASL/NEAMS

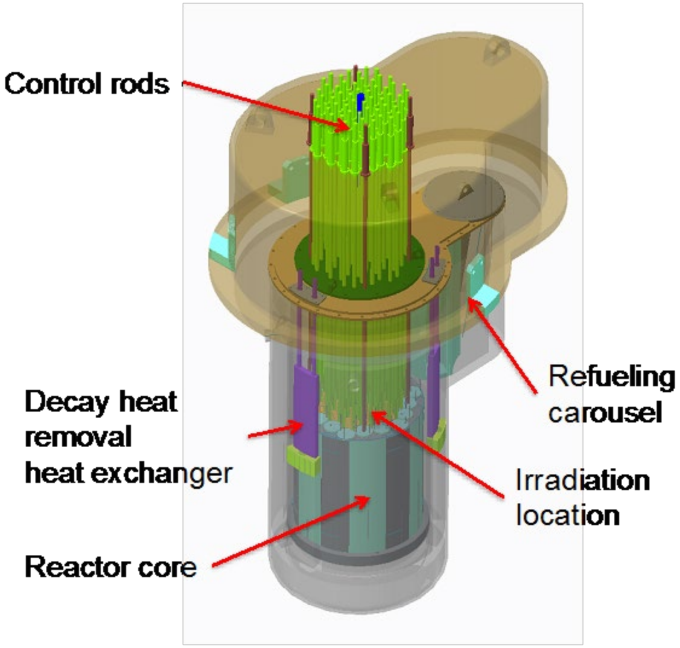
Applications beyond SMRs

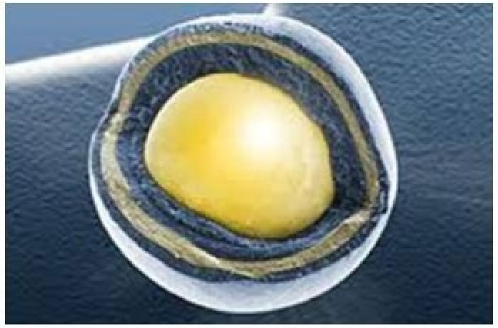


- Advanced reactors – pebble beds, molten salt
- Micro-reactors
- Ex-core vessel fluence and dosimetry
- Radiation shielding








TRISO coated particle fuel