Blowing Up Stars Over the Next Decade

Michael Zingale

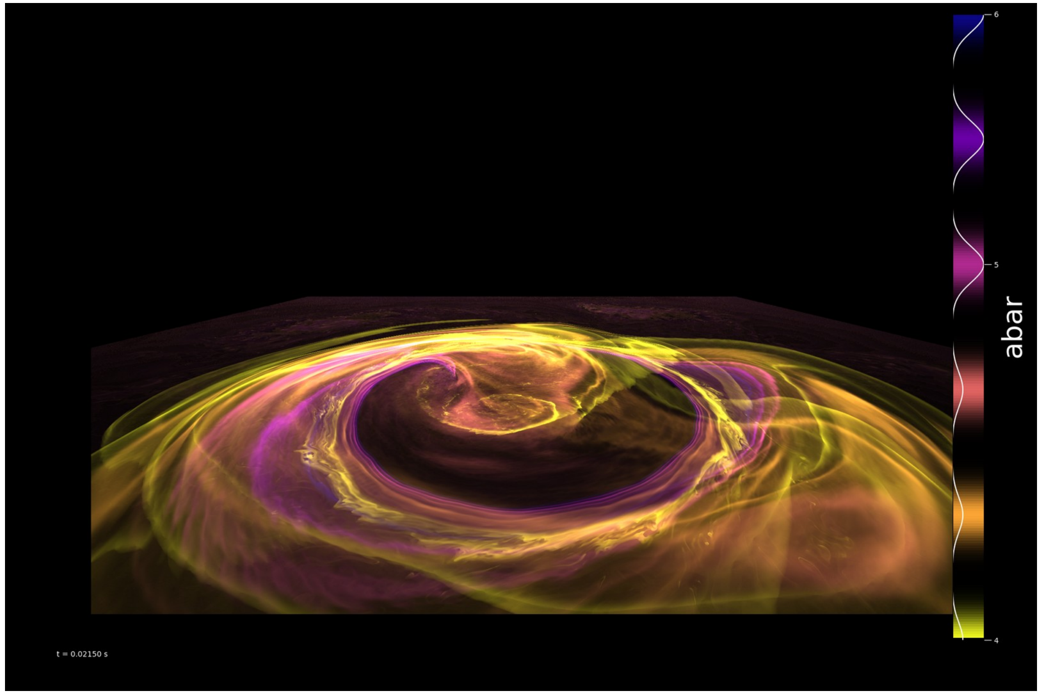
Stony Brook University

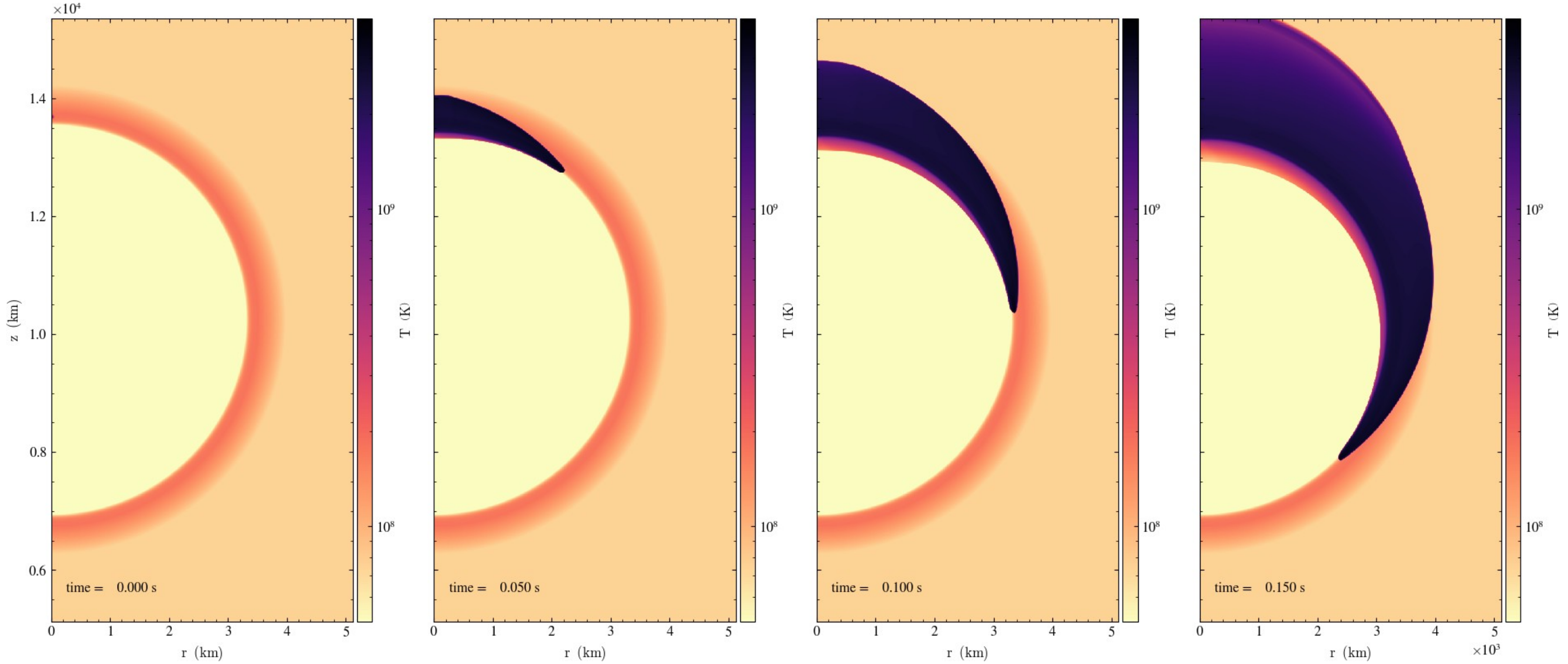
@Michael_Zingale
https://github.com/zingale

Support from DOE Office of Nuclear Physics, DOE ECP, DOE SciDAC
Computer time via DOE INCITE @ OLCF/ORNL and NERSC/LBNL

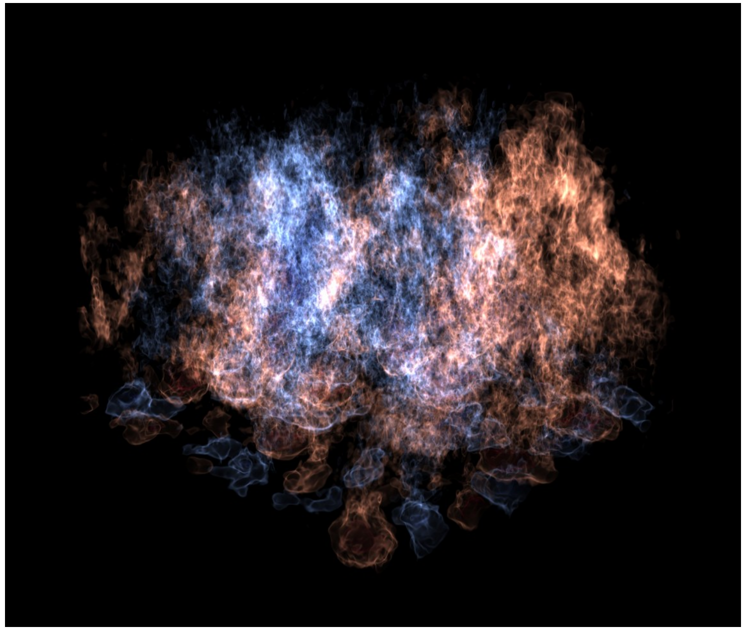
Our Science

- We model stellar explosions
  - Common theme: coupling hydrodynamics and nuclear reactions

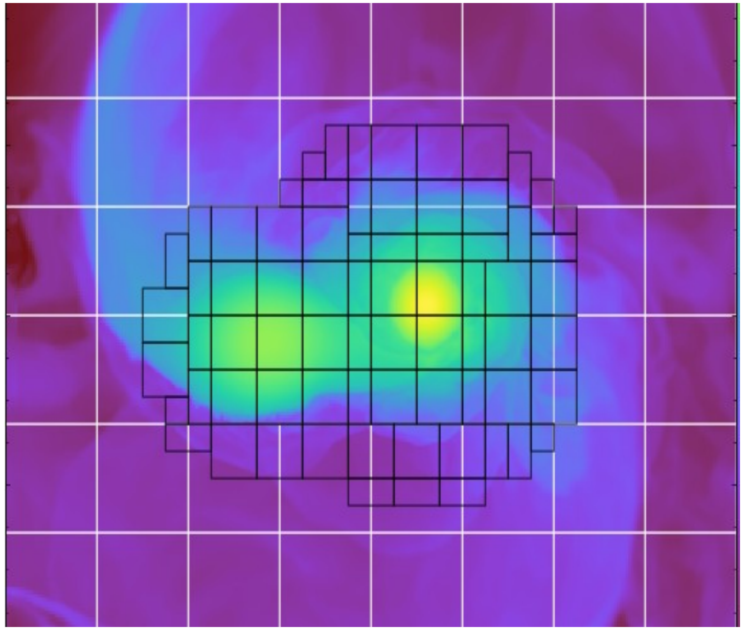




Our Codes

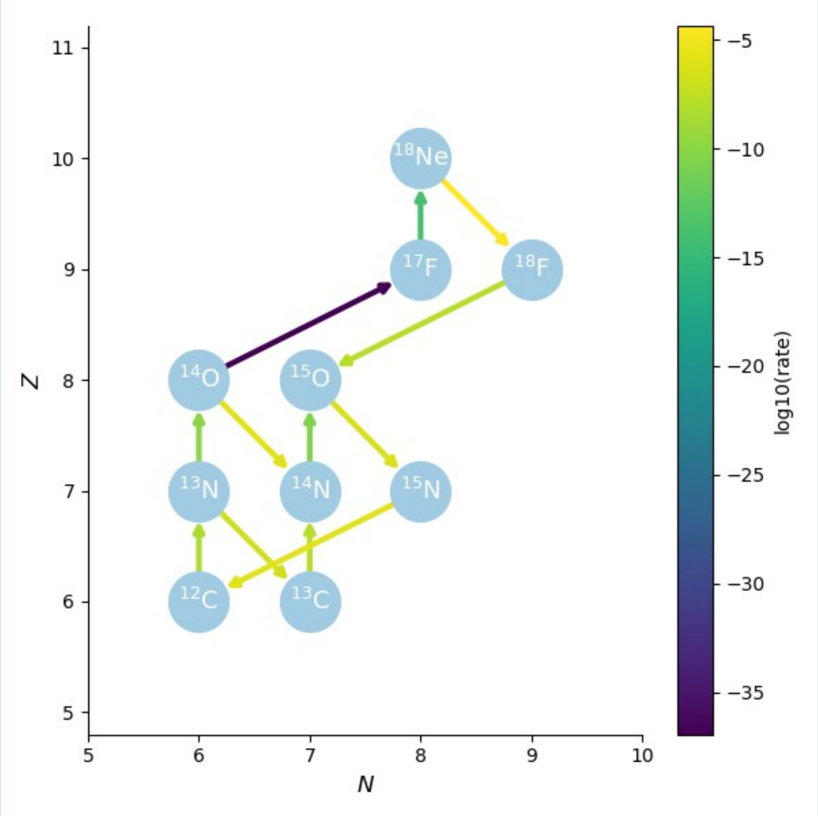


MAESTROeX: low Mach number stratified flows



Castro: compressible (magneto-, radiation-) hydrodynamics

https://github.com/amrex-astro



pynucastro: a python library for nuclear astrophysics

https://github.com/pynucastro

Challenges

- We need to design algorithms that are matched to the science
- Codes continually need to be ported to new architectures
- Computing is becoming more heterogeneous
- Learning curve for new students is increasing greatly
- We all need more computer time

What Did LRP 2015 Say?

- *The use of high-performance supercomputers has revolutionized the modeling of stellar evolution, core collapse supernovae, compact object physics, neutrino astrophysics, and the emergence of structure in the Universe*
- *Multidimensional hydrodynamics calculations are critical for advancing the understanding of stellar explosions, high-energy nuclear collisions, and the detonation of nuclear weapons.*
- *It is essential that the nuclear theory community be prepared to exploit these new architectures, requiring a concerted effort to port scientific codes and to optimize their performance. A significant software development workforce and close collaboration with computer scientists, applied mathematicians, and hardware vendors are required to successfully port our existing scientific code bases, in addition to expanding their scientific reach.*

Access to Computing

- INCITE and ALCC provide access to capability computing
  - Optimized for big “hero jobs”
  - Competition is tight
- What happens if you don’t get time?
  - Grad student projects left in limbo?
  - Need to pivot to other science?

Access to Computing

- NERSC provides capacity computing
  - Most of the time is via ERCAP
- NP allocated ~ 12% of CPU and 16% of GPU node hours
  - Roughly in proportion of Office of Science budgets
  - But HEP fares much better
- Excellent science is done on small scales
  - e.g. parameter studies

DOE NP needs more computer time all around. Pushing for an increase in capacity computing would help the most NP scientists

Total node hours available at NERSC in Allocation Year 2023.

	AY2023 CPU Node Hours Available to Allocate	AY2023 GPU Node Hours Available to Allocate
Advanced Scientific Computing Research (ASCR)	462,100	218,000
Biological and Environmental Research (BER): Earth and Environmental Systems Sciences Division	1,952,300	921,100
Biological and Environmental Research (BER): Biological Systems Science Division	450,500	212,500
Basic Energy Sciences (BES) - Chemical Sciences, Geosciences and Biosciences	2,973,400	1,402,900
Basic Energy Sciences (BES) - Materials Sciences and Engineering	2,540,200	1,198,500
Basic Energy Sciences (BES) - Scientific User Facilities	535,200	252,500
Fusion Energy Sciences (FES)	2,194,900	1,035,600
High Energy Physics (HEP)	2,638,500	1,244,900
Nuclear Physics (NP)	1,963,900	926,600
SBIR	78,500	37,000
Total	**15,789,500**	**7,449,600**

https://www.nersc.gov/users/accounts/allocations/2023-call-for-proposals-to-use-nersc-resources/

New Architectures = New Science

- GPU access has greatly accelerated our work
  - Our XRB simulations run 10 – 20× faster (per node) on GPUs vs CPUs
  - Transforming codes has a cost
    - Spent ~ 1 year porting instead of driving new science

- GPUs can be a poor match
  - Reaction nets and particle Monte Carlo have trouble
    - Lots of memory per zone / particle = register pressure
    - Extreme variation in work across zones / particles

HPC is not driving new GPU / processor tech. DOE will need to pay a premium for hardware matched to science priorities

Will AI Leave Us Out?

- Several NP groups are exploring how machine learning can help
  - A lot of GPU development is focused on AI features
  - We've explored using neutral networks for reaction ODEs
- Next generation of machines may put even more emphasis of machine learning
- We lack people trained in both AI/ML and science

DOE should staff computing centers with AI/ML experts to act as liaisons to the science teams.

Algorithm / Software Dev Happens in NP

- As domain scientists, we know best where the “trouble spots” are in our codes
- Examples from our group:
  - New coupling methods for hydrodynamics + reactions based on spectral deferred corrections
  - pynucastro: python library for nuclear astrophysics
- Metrics don’t tell the whole story
  - It can take years for a code development to start producing science
  - Simply looking at cites in the last X years misses the story

Grant reviews need to recognize that code development operates on a different timescale than theory

Continued Training is Essential

- New students / postdocs need to come up to speed on
  - Algorithms used in our field
  - MPI parallelism
  - GPU offloading
  - Emerging technologies (e.g. AI/ML)
- Current training events work well:
  - ATPESC
  - Hackathons
- Need to grow opportunities in the future

Training opportunities for NP students will become even more important over the next decade as computing technologies explode

NP Should Require Open Source

- Science requires that we share our methodology
- Simulation codes are too complex to fully describe in papers
- Code should be available at the same time as papers are published

Just like the Open Access publication requirement, DOE should require code to be deposited in a DOI repo (like Zenodo) when a paper is published.

Supporting Users is Important

- Our simulation codes are tools are used throughout the community
- Grants should recognize that supporting users is a cost

Metrics for reviewing grants should include code contributions to community.