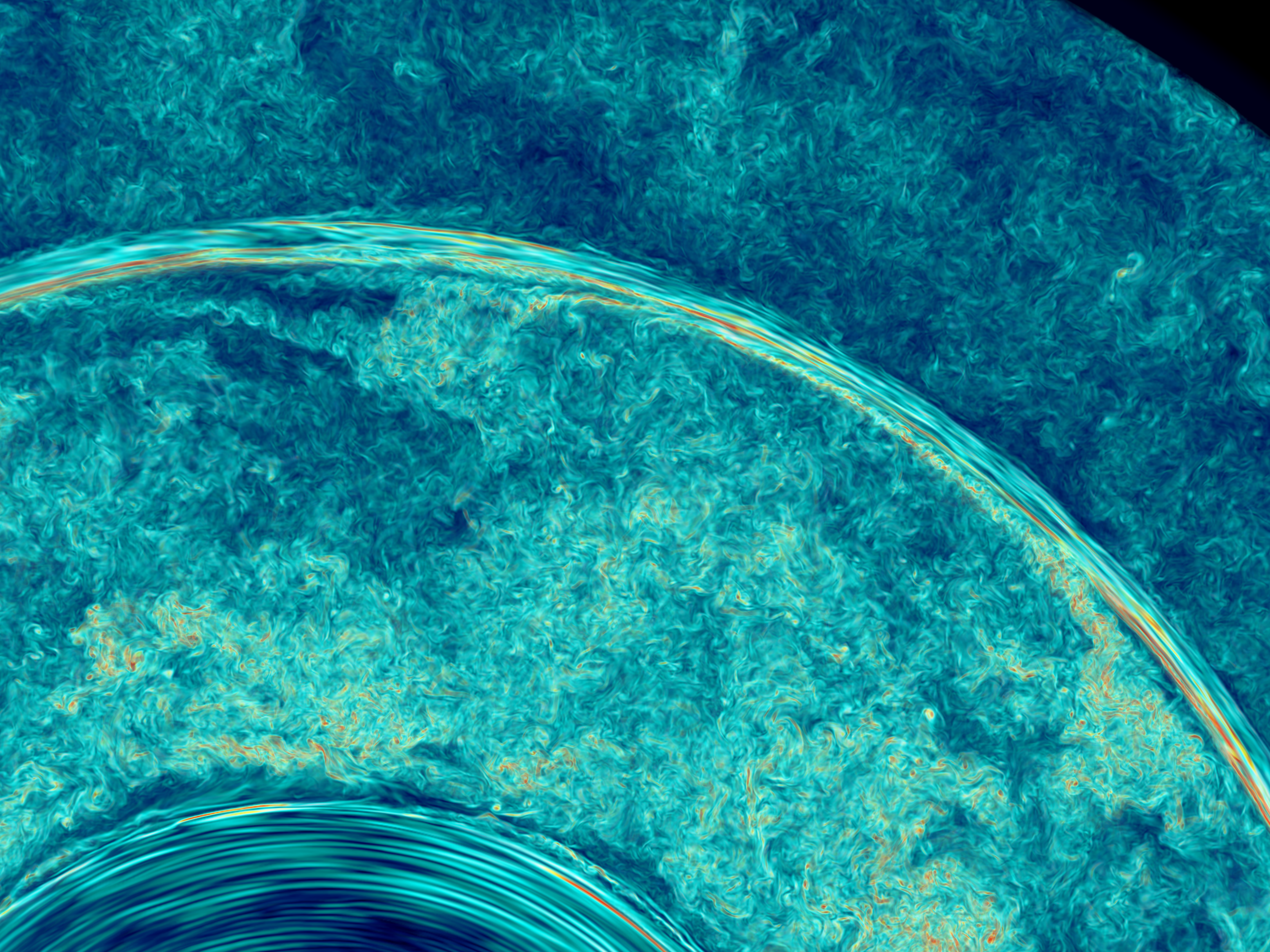
Case study: Large-scale 3D simulations of convection in a star

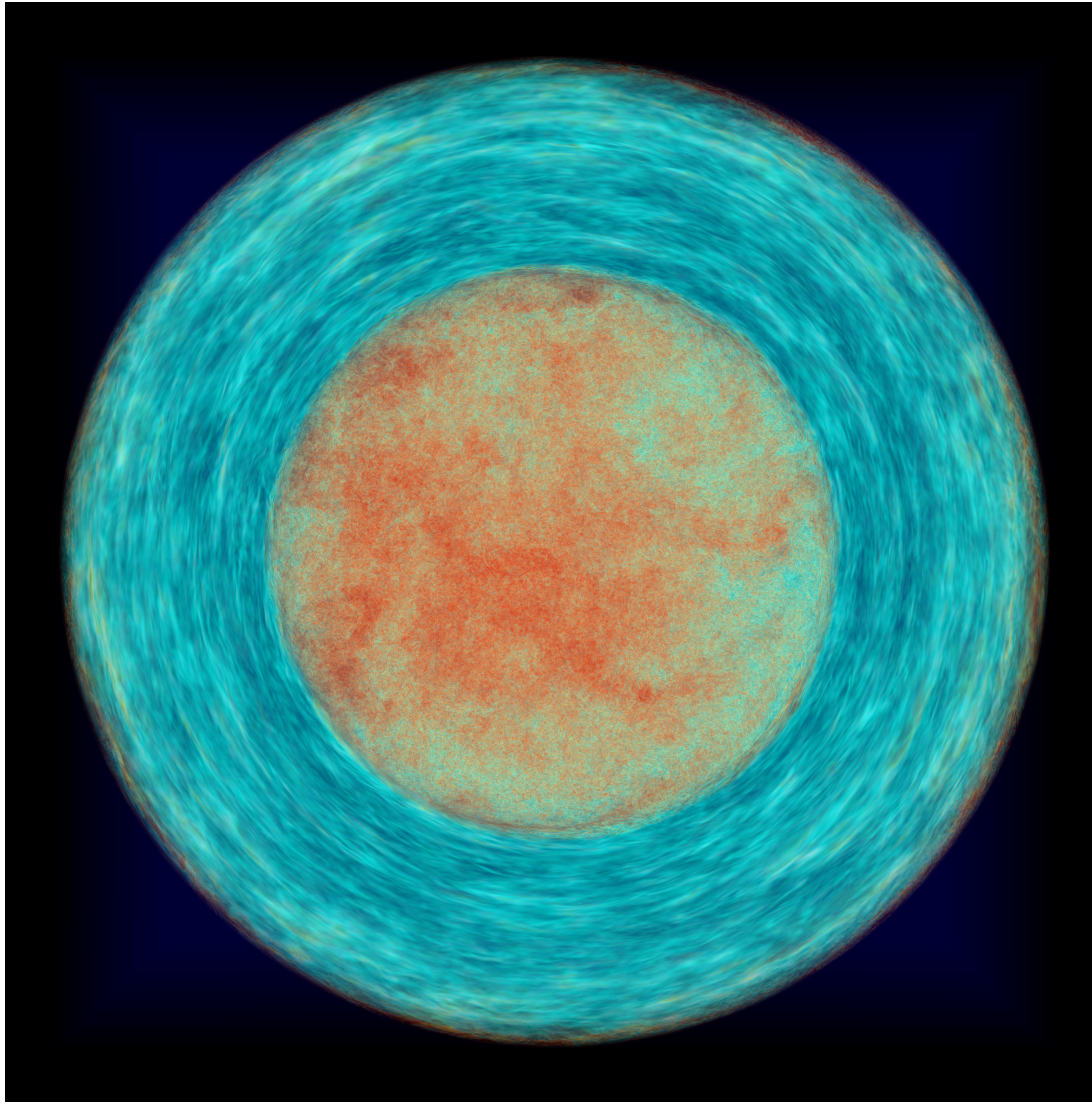
https://www.ppmstar.org virtual research environment



Volume rendering of magnitude of vorticity of a central slice of a 4π full sphere simulation of $25M_{\odot}$ H-core burning convection.

The highly turbulent nature of stellar convection is clear in these simulations from the irregular yet locally isotropic distribution of vorticity.

Simulation performed with the PPMstar code (Paul Woodward, University of Minnesota)



Large parallel runs on Niagara

https://docs.computecanada.ca/wiki/Niagara

compute canada calcul canada

Page | Discussion | Read | View s

Wiki Main Page

Support

- Getting started
- Getting help
- Running jobs
- Known issues
- System status

Resources

Niagara

Other languages:	English • français

Availability: In production since April 2018	
Login node: **niagara.computecanada.ca**	
Globus endpoint: **computecanada#niagara**	
Data mover nodes (rsync, scp, ...): **nia-dm2, nia-dm2**, see Moving data	
System Status Page: **https://docs.scinet.utoronto.ca**	

Niagara hardware specifications [edit]

- 1548 nodes, each with 40 Intel Skylake cores at 2.4GHz, for a total of 61,920 cores.
- 202 GB (188 GiB) of RAM per node.
- EDR Infiniband network in a so-called 'Dragonfly+' topology.
- 7PB of scratch, 2PB of project space (parallel filesystem: IBM Spectrum Scale, formerly known as GPFS).
- 256 TB burst buffer (Excelero + IBM Spectrum Scale).
- No local disks.
- No GPUs.
- Theoretical peak performance ("Rpeak") of 4.75 PF.
- Measured delivered performance ("Rmax") of 3.07 PF.
- 685 kW power consumption.

About effective large-scale parallel simulations

- Key challenge: processing on cores is fast but getting data there is slow
- Overlap of communication and computation
- Set aside team leaders to deal with communications while workers continue to work
- Compress messages into smaller number of large messages, requires finding balance, depends on hardware
- Processors get faster but the network is not getting faster at the same rate

- Dedicated code to do just this type of problem, not general purpose code
- $1.2^{10}=6.2$ → pick up each 10-20% effect, for example
- Beware of the different flop cost of different operations, there is a logical operation for addition or multiplication (1 flop each) but not for sin, cos or divide (14 flops each)
- Solve problems the way computers like them, often not how mathematicians would do it

References:

Woodward PR, Herwig F, Wetherbee T. Comput. Sci. Eng. 20(5):8–17
Woodward PR, Lin P-H, Mao H, Andrassy R, Herwig F. 2018. eprint arXiv:1810.13416 (ASTRONUM conference paper)

Cyber challenges

- Data deluge naive approach: one dump 4 quantities 1536^3 x 8 byte x 4 quantities = 115GB - disk read (100MB/s) 20min, network transport (20MB/s) 1.6hrs
- 500 time steps = 57.5TB - disk read 6.6d, network transport 33d
- Large data sets require remote access and smart data management strategies

- Multi-scale, multi-physics, multi-method
- Complex data interactions and collaborative cyber- research environments, data fusion

- Research data management - in reality means combining access to data, data-specific analytics tools and capability to execute analytic tools to data
- Sharing work flows
- Legacy software, tools, and workflows
- Reproducibility of science

Data management for PPMstar code

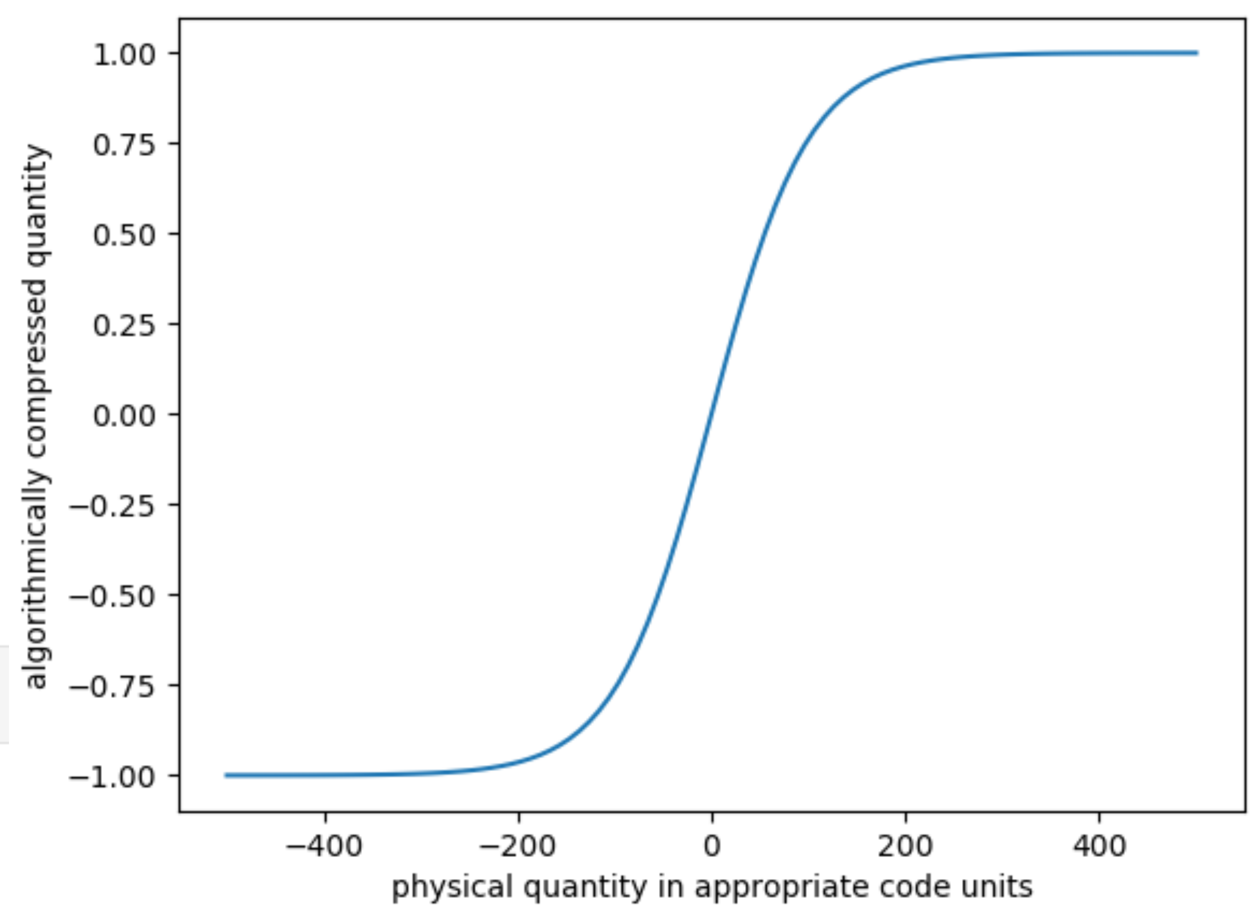
- In-code generation of science-ready analysis data and images (this is a substantial departure from common practice)
  - instead of just writing out *raw* data we generate three types of algorithmically compressed file types
  - bobs - bricks of bytes algorithmically compressed
  - briquette data - spatially filtered (by factor 4 in each direction, 4 bytes, but 32 primary and derived quantities)
  - rprofs - radial profiles of lots of things
  - as well as one image for each of 10 quantities for each output dump
  - New feature: data compression (zip) takes place in code and data is written algorithmically and data compressed as well as filtered

- Example: one 768^3-grid run with 1M time steps writes out 466 dumps
  - data held on project for immediate analysis (briquette & rprofs): 260GB
  - data on nearline: 3 restart dumps and bobs of 10 quantities (631GB)
  - instead of:

```
d = 466 *8*10* 768**3 /1.e12
print("Uncompressed data volume: 466 * 10 * 8 * 768**3 /1.e12 =  {:.2f}TB".format(d))
```

```
Uncompressed data volume: 466 * 10 * 8 * 768**3 /1.e12 =  16.89TB
```

Example for algorithmic data compression



Sharing computational science with the community

Verification: Do you solve the equations right?

Validation: Do you solve the right equation?

Andrassy+ 22, A&A

Dynamics in a stellar convective layer and at its boundary: Comparison of five 3D hydrodynamics codes

R. Andrassy[1], J. Higl[1], H. Mao[2], M. Mocák[3], D. G. Vlaykov[4], W. D. Arnett[5], I. Baraffe[4,6], S. W. Campbell[7,8], T. Constantino[4], P. V. F. Edelmann[9], T. Goffrey[10], T. Guillet[4], F. Herwig[11,12], R. Hirschi[3,13], L. Horst[1], G. Leidi[1,14], C. Meakin[5,15], J. Pratt[16], F. Rizzuti[3], F. K. Röpke[1,17], and P. Woodward[2,12]

R. Andrassy et al.: Comparison of five 3D hydrodynamics codes

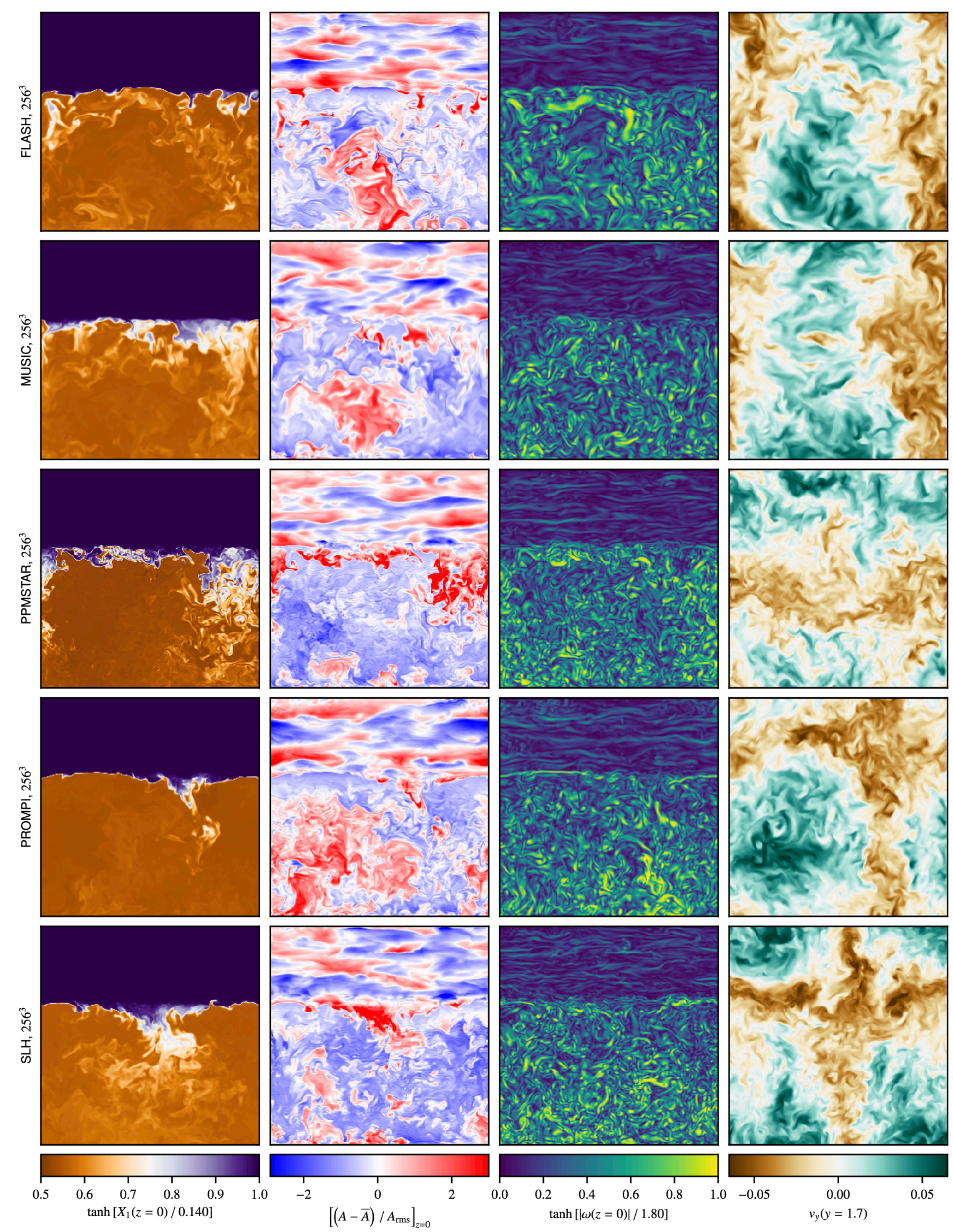
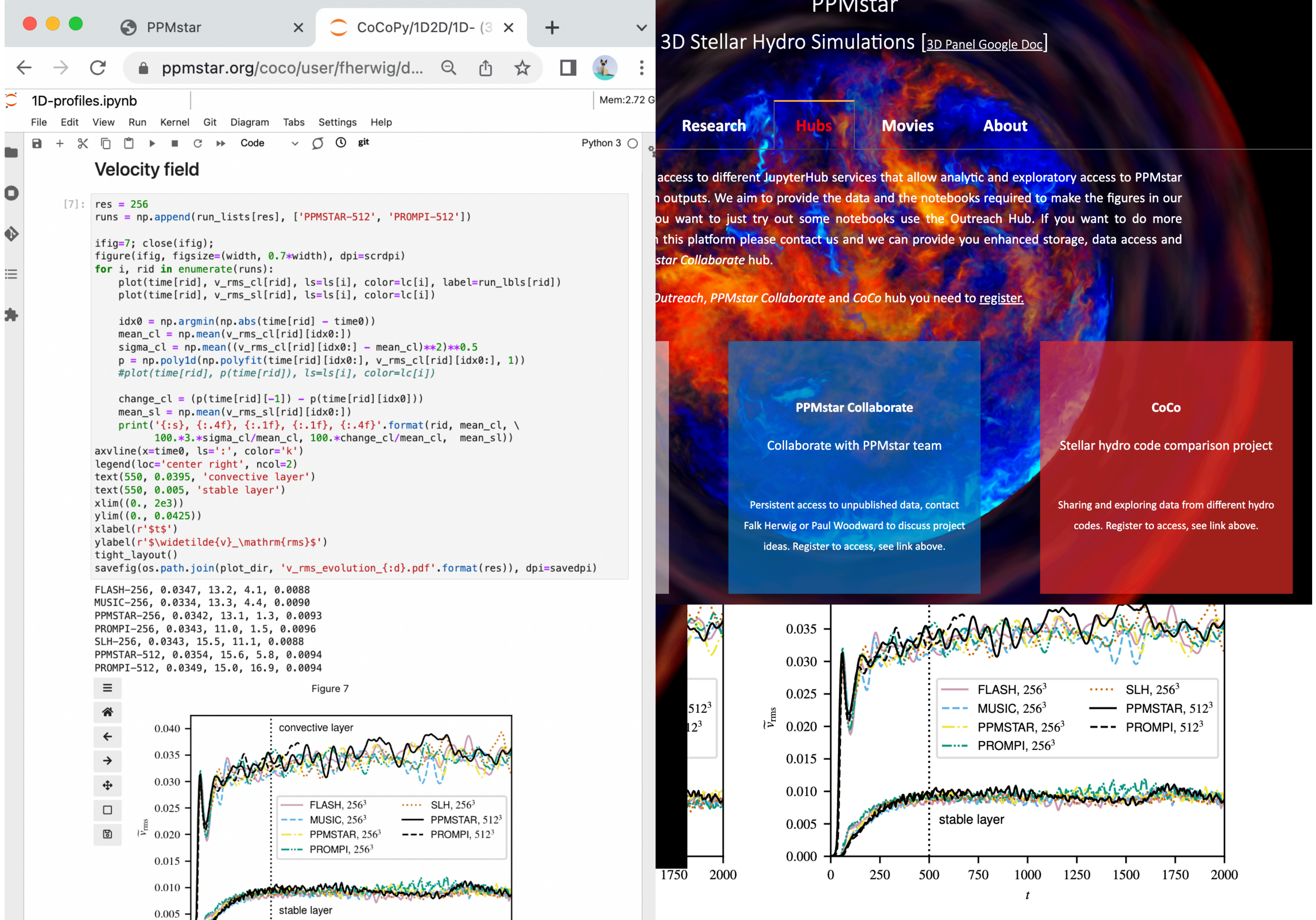


Fig. 2. Renderings of the flow field at $t = 1000$. The variables shown (from left to right) are the mass fraction X_1, relative entropy fluctuations $(A - \overline{A})/A_{\rm rms}$, magnitude of vorticity $|\omega|$, and the vertical component of velocity v_y. The first three variables correspond to slices through the simulations box in the $z = 0$ plane and the last one in the $y = 1.7$ plane. The non-linear scaling of X_1 and $|\omega|$ and the normalisation of entropy fluctuations by their horizontal rms spread $A_{\rm rms}$ (see the colour bars) are introduced for visualisation purposes.



velocity fluctuations in the convection zone ($y < y_{\mathrm{ub}}(t) - 0.1$, upper set of curves) and
he location $y_{\mathrm{ub}}(t)$ of the convective boundary is tracked in time as shown in Fig. 9. The
cluded from the analysis.