

Profiles of emerging HPC applications

Scientific Achievement

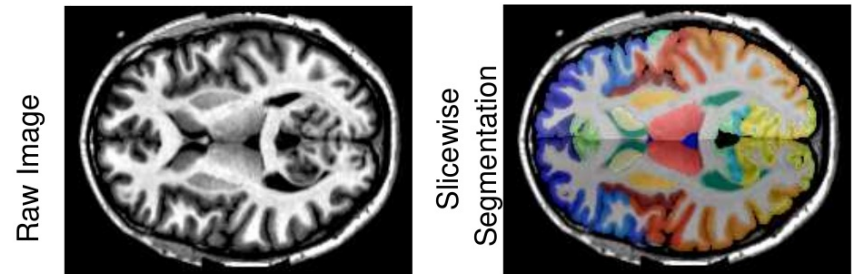
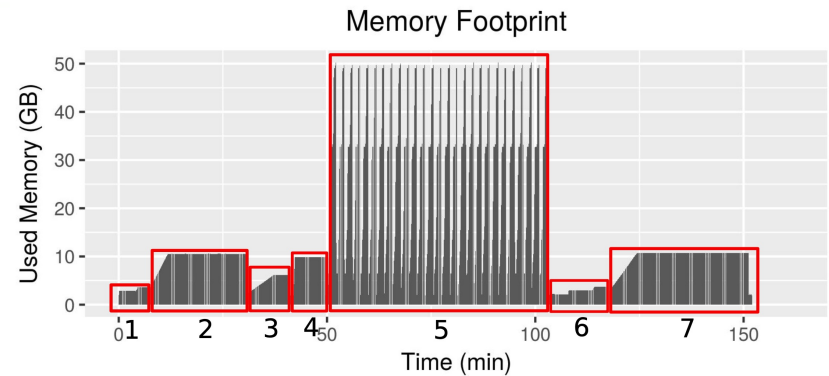
Analytical model for describing the behavior of emerging HPC applications from fields like neuroscience, genome research and bioinformatics that include highly stochastic and dynamic resource requirements for both CPU and memory that cannot easily be predicted before runtime.

Significance and Impact

The model has been applied in a scheduling framework to allow applications that do not follow the traditional HPC model to be deployed efficiently on large-scale systems and not rely on ad-hoc inefficient solutions for predicting their resource requirements.

Research Details

- The model describes applications as a chain of tasks each with probabilistic resource requirements based on as little as 10 runs and can adapt to shifts in behavior.
- Applications from fields doing exploratory research that do not fit the HPC model can optimize their submissions and request more realistic resources.
- Experiments for applications from the neuroscience department at the Vanderbilt university show an average improvement of 20-25% in response time.



Example of a brain segmentation application that can be deployed more efficiently when modeling its walltime and memory footprint

A. Gainaru, B. Goglin, V. Honoré and G. Pallez,
Profiles of Upcoming HPC Applications and Their Impact on Reservation Strategies in *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 5, pp. 1178-1190
doi: <https://doi.org/10.1109/TPDS.2020.3039728>



U.S. DEPARTMENT OF
ENERGY

Office of
Science



VANDERBILT
UNIVERSITY

Inria

OAK RIDGE
National Laboratory