PFASST with dynamic resource management for large-scale applications

Dominik Huber, Hans-Joachim Bungartz, Jan Fecht, Keerthi Gaddameedi, Martin Schreiber, Martin Schulz, Tobias Neckel, Valentina Schüller

Abstract

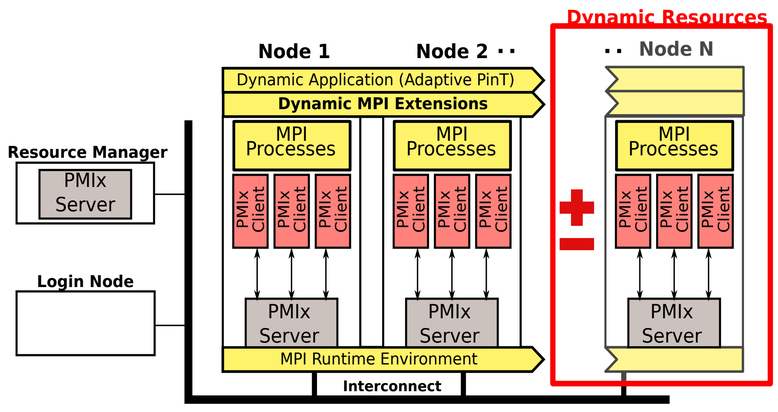
Dynamic resource management is a fairly recent development made to improve the flexibility of job scheduling.

MPI-based approach [1] makes this possible by treating resources on a set granularity and performs set operations whenever changes are made. Due to the extreme resource requirements, using dynamic resources will likely be compulsory for parallel-in-time methods.

Development towards this has been made by prototyping an interface for dynamic resource management and then incorporating it into LibPFASST to provide proof of concept. The intent of this poster is to motivate further investigation and exploration of dynamic resources for large-scale applications such as weather/climate simulations, molecular dynamics, graph design and more.

Dynamic Resource Management

Prototype implemented on top of OpenMPI, OpenPMIx and PRRTE



Provides the ability to add and remove resources assigned to a job dynamically during program execution.

Main Benefits:

- Utilization rate
- Job throughput
- Energy efficiency
- Program efficiency
- Fault tolerance

Design Principles:

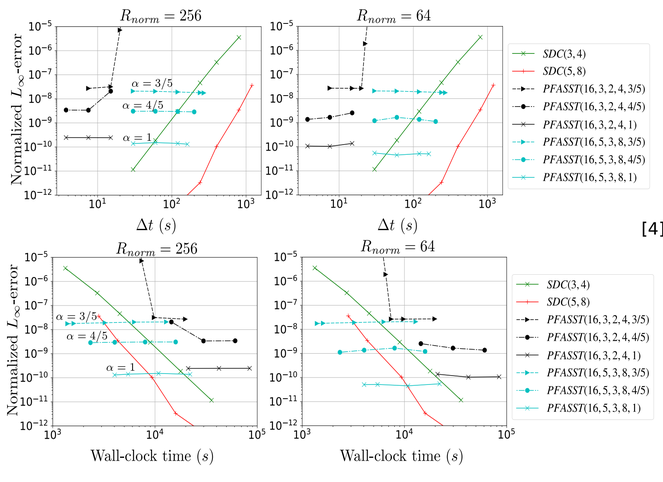
1. Granularity of resource changes → Sets of resources.
2. Resource changes → Set operations.
3. Resource sets information → Associated data store.
4. Indepedent of software stack specifics → Generic.
5. Cooperate between applications and the resource manager.

Extension of MPI Sessions & PMIx interface to enable dynamic resources for PMIx- and/or MPI-based applications.

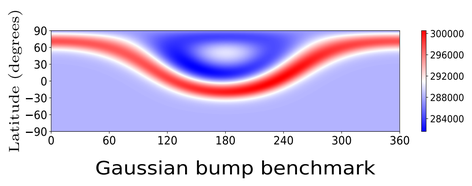
Integration of SWEET with LibPFASST

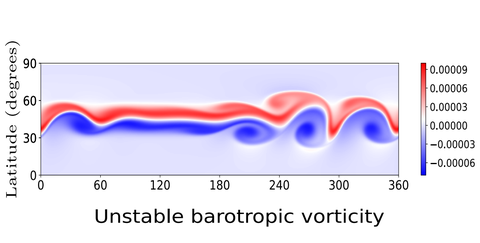
SWEET - **S**hallow **W**ater **E**quation **E**nvironment for **T**ests

A C++ solver that allows investigation and prototyping of various numerical methods for dynamical cores of atmospheric models.



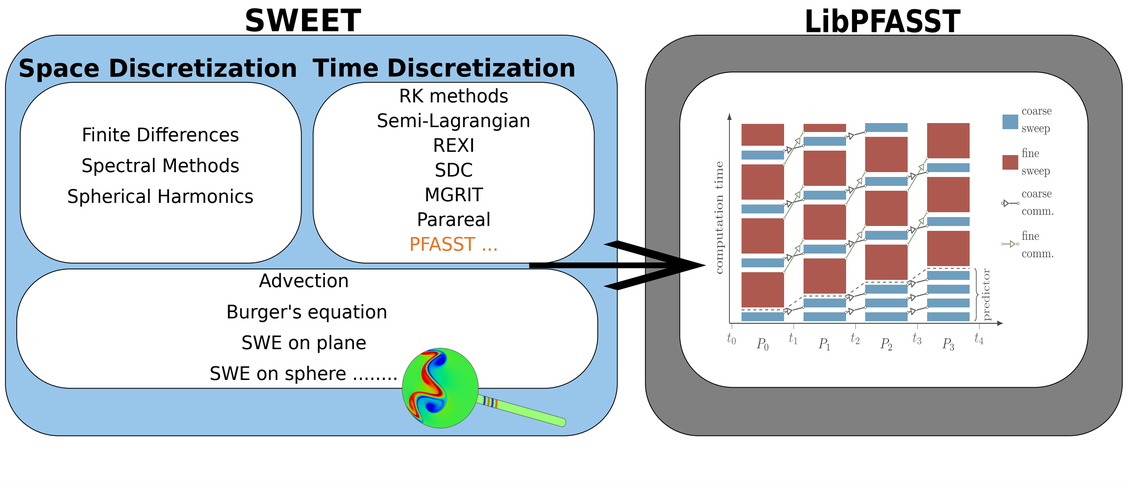
[4]





Coupling objective

To test the performance of PFASST in the context of atmospheric simulations and impact of various parameters.



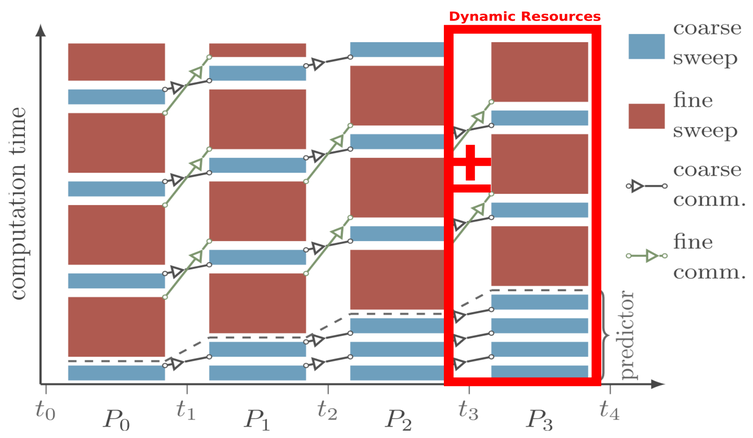
LibPFASST with Dynamic Resources

LibPFASST - **P**arallel **F**ull Approximation **S**cheme in **S**pace and **T**ime

A fortran implemention of PFASST [3], a high order time integration scheme that can be viewed as a parareal version of MLSDC [5].

Motivation

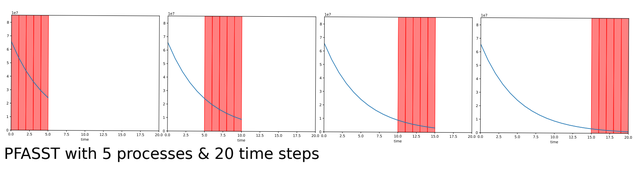
1. Increasing/decreasing the number of parallel-in-time steps in PinT methods is uncomplicated.
2. Convergence behaviour is problem-dependent and optimal parallel-in-time steps may vary.



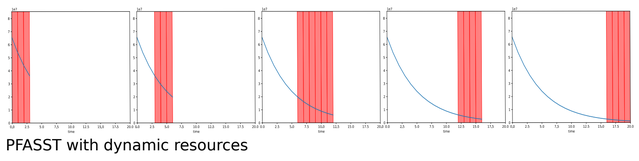
Modified LibPFASST to support dynamic resource changes on top of modified MPI runtime.

Resizing hook informs library on resource adjustments between PinT blocks.

Resources are adapted based on convergence behavior and other metrics (outlook).

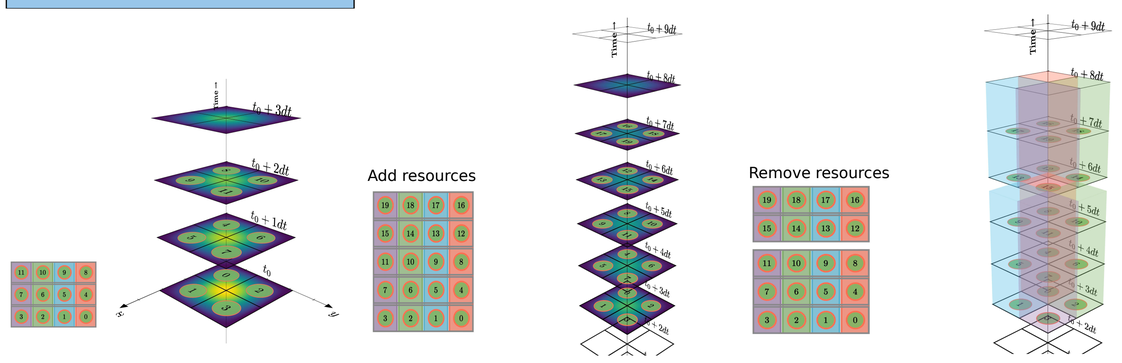


PFASST with 5 processes & 20 time steps



PFASST with dynamic resources

PoC: 2D Heat equation.
12 processses, 3 parallel-in-time steps



Future work:

1. Incorporate dynamic resource management into SWEET.
2. Investigate this concept with shallow water equations.
3. Expand investigation into other large-scale applications.

DynRes + **LibPFASST** + **SWEET**

References

[1] Huber, D., Streubel, M., Comprés, I., Schulz, M., Schreiber, M., Pritchard,HH.: Towards dynamic resource management with mpi sessions and pmix. In:Proceedings of the 29th European MPI Users' Group Meeting. p. 5767. Eu-HroMPI/USA'22, Association for Computing Machinery, New York, NY, USA(2022). https://doi.org/10.1145/3555819.3555856.

[2] Fecht, J., Schreiber, M., Schulz, M., Pritchard, H., Holmes, D.J.: An Emulation Layer for Dynamic Resources with MPI Sessions. In: HPCMALL 2022 - Malleability Techniques Applications in High-Performance Computing. Hambourg, Germany (Jun 2022), https://hal.science/hal-03856702.

[3] M. Emmett and M. Minion, “Toward an efficient parallel in time methodfor partial differential equations,” Communications in Applied Mathemat-ics and Computational Science, vol. 7, no. 1, pp. 105–132, Mar. 28, 2012,issn: 2157-5452, 1559-3940. doi: 10.2140/camcos.2012.7.105. [Online].Available: http://msp.org/camcos/2012/7- 1/p04.xhtml (visited on03/16/2021).

[4] F. P. Hamon, M. Schreiber, and M. L. Minion, “Parallel-in-Time Multi-Level Integration of the Shallow-Water Equations on the Rotating Sphere,”Journal of Computational Physics, vol. 407, p. 109 210, Apr. 2020, issn:00219991. doi: 10 . 1016 / j . jcp . 2019 . 109210. arXiv: 1904 . 05988.[Online]. Available: http : / / arxiv . org / abs / 1904 . 05988 (visited on03/07/2022).

[5] R. Speck, D. Ruprecht, M. Emmett, M. Minion, M. Bolten, and R. Krause,“A multi-level spectral deferred correction method,” BIT Numerical Math-ematics, vol. 55, no. 3, pp. 843–867, Sep. 2015, issn: 0006-3835, 1572-9125.doi: 10.1007/s10543-014-0517-x. arXiv: 1307.1312. [Online]. Available:http://arxiv.org/abs/1307.1312 (visited on 03/16/2021).

SWEET

dynres

LibPFASST with dynres support

This project has received funding from the Federal Ministry of Education and Research and the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955701, Time-X. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and Belgium, France, Germany, Switzerland.