Goldschmidt 2025 Abstract
https://doi.org/10.7185/gold2025.30854

Accelerating parallel reactive transport simulations by interpolation from automatically clustered Distributed Hash Tables

MARCO DE LUCIA[1], MAX LÜBKE[2] AND BETTINA SCHNOR[2]

[1]GFZ Helmholtz Centre for Geosciences
[2]University of Potsdam

Reactive transport models are computationally very intensive, the bottleneck being usually represented by goechemistry. They therefore require hardware resources which are not always readily available to the geoscientists. Approximated computing, including the use of surrogate models based on machine learning and artificial intelligence (ML/AI), represents a possible solution to this problem. We present here an algorithm implemented in the massively parallel POET scientific simulator [1] in which the input-output values from all calls to the geochemical solver during coupled simulations are stored in distributed Distributed Hash Tables and automatically clustered using a so called *hashing trick*. This allows to efficiently check if similar simulations were previously run and reuse them as basis for interpolating new results. We evaluate the efficiency gain from this technique based on a 2D mock-up model of a microfluidic experiment inducing barite precipitation at the interface of celestite disks. The reference model runs in 10 hours on 96 CPUs and enabling interpolation allowed to cut in half either the computing time or the hardware resources, with no significant accuracy loss, but in either case significant less energy consumption. The relative gain due to interpolation increases with fewer CPUs are employed. Another advantage of this approach is that it does not require any pre-training.

[1] De Lucia, Kühn, Lindemann, Lübke & Schnor (2021), *Geoscientific Model Development* 14, 7391--7409. https://doi.org/10.5194/gmd-14-7391-2021