Performance Evaluation of in situ Applications through Simulation using SimGrid

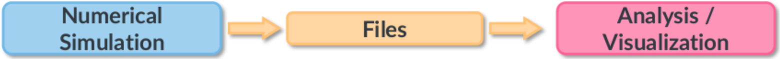
Valentin Honoré

ENSIIE & Samovar, Evry - Palaiseau

Presentation of HPDA/PDS Research Projects

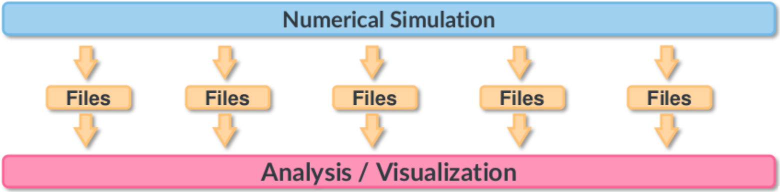
In situ processing: historical meaning

- Post-hoc

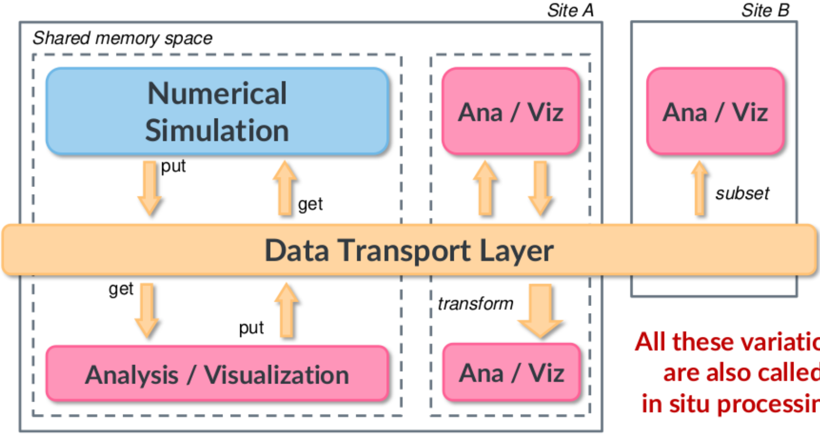


☹ Files becoming too big + Gap increase between CPU and I/O ☹

- In situ



In situ processing: modern approach



All these variations are also called in situ processing!

- How much resources give to simulation, analysis, viz? [ALLOCATION]
- Where and when run the analysis/visualization? [MAPPING]
- Files out of question? In-memory or network? [DATA TRANSPORT]

Sim-Situ: a tool for performance evaluation using Simgrid

- ▷ Answer RQs → **Performance evaluation**
  - Take good decisions → **Objective performance indicators**
- ▷ **Go beyond the traditional empiric guess** → **Speed and Flexibility**
  - Explore many unconventional scenarios
  - Consider unconventional performance metrics

Experiments	**Simulation**
→ Time- and resource-consuming	→ Run on a laptop
→ Complex to set up	→ Highly flexible
→ Limited in scope	→ What-if scenarios
→ Sensitive to exogenous factors	→ Reproducibility and control

6

Goal of the project

- **Goal:** enhance the Sim-Situ tool introduced in [1]
- Proposed steps of the project
  1. Create a Git repository with Continuous Integration for Sim-Situ [1]
  2. Test Sim-Situ with different applications: other proxy-applications such as CabanaMD or QuickSilver, more complex simulation codes (such as Gromacs [2], LAMMPS [3] etc)
  3. Code refinement: introduce generic descriptions of simulation and analysis components
  4. Externalize the application description: description of analysis/viz component
  5. Add new features: evaluation of more mapping and allocation strategies, implementing state-of-the-art DTLs (ADIOS [4], DataSpaces [5]) etc
- Requirements: Git, taste for programming and compiling

V. Honoré, T. M. A. Do, L. Pottier, R. Ferreira da Silva, E. Deelman, and F. Suter, "SIM-SITU: A Framework for the Faithful Simulation of in situ Processing," in eScience 2022, Salt Lake City, United States, Oct. 2022. [Online]. Available: https://hal.archives-ouvertes.fr/hal-03504863

P. Bauer, B. Hess, and E. Lindahl, "Gromacs 2022.3 manual," Tech. Rep., Sep. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.7037337

S. Plimpton, "Fast Parallel Algorithms for Short-Range Molecular Dynamics," Journal of Computational Physics, vol. 117, no. 1, pp. 1–19, 1995.

W. F. Godoy, N. Podhorszki, R. Wang, C. Atkins, G. Eisenhauer, J. Gu, P. Davis, J. Choi, K. Germaschewski, K. Huck, A. Huebl, M. Kim, J. Kress, T. Kurc, Q. Liu, J. Logan, K. Mehta, G. Ostrouchov, M. Parashar, F. Poeschel, D. Pugmire, E. Suchyta, K. Takahashi, N. Thompson, S. Tsutsumi, L. Wan, M. Wolf, K. Wu, and S. Klasky, "ADIOS 2: The Adaptable Input Output System. A framework for high-performance data management," SoftwareX, vol. 12, p. 100561, 2020.

C. Docan, M. Parashar, and S. Klasky, "DataSpaces: an Interaction and Coordination Framework for Coupled Simulation Workflows," Cluster Computing, vol. 15, no. 2, pp. 163–181, 2012.