Contribution ID: **46**   Type: **Student Micro-talk**

Composition of an Earth Science Modeling Workflow in Hybrid HPC

SOMOSPIE [1] is a modular workflow that predicts fine-grained soil moisture from coarse-grained satellite data and terrain parameters. Soil moisture is important for environmental sciences, precise agriculture, and wildfire detection. Depending on the resolution, the data of SOMOSPIE for a region of interest such as Oklahoma (a rich agriculture area) can range from 4.5 GB to 42 TB. The large amount of data requires an elastic, flexible but high-performance computing platform to support the workflow execution.

Hybrid HPC (HPC on Cloud) offers the security and powerful performance of HPC and the flexibility and scalability of the cloud. As a consequence, HPC on the cloud is becoming an alternative to on-premise clusters for executing scientific applications [2][3].
We leverage the advantages of Hybrid HPC and compose SOMOSPIE in the IBM HPC Cluster (IBM Spectrum LSF). Furthermore, we define the best practices of composing an HPC application in the Cloud in terms of data and workflow orchestration, storage, dynamic composition, and network.

As a result of this work, SOMOSPIE can be deployed with multiple and separate scenarios of data resolutions on customized architectures on the IBM HPC Cluster and portable platforms through container technology such as Docker [4] and Singularity [5].

References

[1] D. Rorabaugh, M. Guevara, R. Llamas, J. Kitson, R. Vargas and M. Taufer, "SOMOSPIE: A Modular SOil MOisture SPatial Inference Engine Based on Data-Driven Decisions," 2019 15th International Conference on eScience (eScience), 2019, pp. 1-10, doi: 10.1109/eScience.2019.00008.
[2] D. Sokolowski, J. -P. Lehr, C. Bischof and G. Salvaneschi, "Leveraging Hybrid Cloud HPC with Multitier Reactive Programming," 2020 IEEE/ACM International Workshop on Interoperability of Supercomputing and Cloud Technologies (SuperCompCloud), 2020, pp. 27-32, doi: 10.1109/SuperCompCloud51944.2020.00010.
[3] Marco A. S. Netto, Rodrigo N. Calheiros, Eduardo R. Rodrigues, Renato L. F. Cunha, and Rajkumar Buyya. 2018. HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. ACM Comput. Surv. 51, 1, Article 8 (April 2018), 29 pages. DOI:https://doi.org/10.1145/3150224
[4] Merkel, D. (2014). Docker: lightweight Linux containers for consistent development and deployment. Linux Journal, 2014(239), 2.
[5] Kurtzer GM, Sochat V, Bauer MW (2017) Singularity: Scientific containers for mobility of compute. PLOS ONE 12(5): e0177459. https://doi.org/10.1371/journal.pone.0177459

Student?

Yes

Supervisor name

Michela Taufer

Supervisor email

taufer@acm.org

Primary authors: OLAYA, Paula; Mrs TAUFER, Michela (The University of Tennessee Knoxville); LOF-STEAD, Jay (Sandia National Laboratories); WEN, Sophia (IBM); CHUNG, I-Hsin (IBM); SEELAM, Seetharami (IBM)

Presenter: OLAYA, Paula

Session Classification: Micro-talks

Track Classification: Earth Systems Modelling