Extreme Data Processing in Exascale Systems

Dana Petcu

West University of Timişoara, Romania

Extended abstract

Extreme data reffers to massive amounts of data that must be queried, communicated and analyzed in near real-time. Analyzing massive scientific data gathered in each second, mining huge sets of images in crysis management, or dealing with millions of social data posts for are few examples. Processing them on exascale systems requires new programming models [1]. Data-aware basic operations for data-intensive applucations supporting the scale use of a massive number of processing elements were introduced recently [2] (a bird-view of these proposals will be presented in the talk). As follow ups, solutions for anomaly and event detections [3, 4] and monitoring in data processing [5] were proposed (a focused view of these proposal will be provided in the talk). Howewer, moving part of the data (pre-)proceesing out of the exascale system towards the edge of the network is another solution and techniques like transprecision computation or machine learning to adapt to the device capabilities are able to reduce the load for the exascale system (particular solutions proposed in [6, 7, 8] will be reviewed in the talk).

References

[1] G. Da Costa, T. Fahringer, J.-A. Rico-Gallego, I. Grasso, A. Hristov, H.D. Karatza, A. Lastovetsky, F. Marozzo, D. Petcu, G. L. Stavrinides, D. Talia, P. Trunfio, H. Astsatryan, Exascale Machines Require New Programming Paradigms and Runtimes, Supercomputing Frontiers and Innovations vol 2, No. 2, 6-27, 2015

[2] ASPIDE Consortium, Exascale programIng models for extreme data processing, European Unions Horizon 2020 research and innovation programme under grant agreement No 801091, https://www.aspide-project.eu/publications/

[3] G. Iuhasz, D. Petcu, Perspectives on Anomaly and Event Detection in Exascale Systems, Proceedings of 2019 IEEE International Conference on High Performance and Smart Computing (HPSC 2019), Washighton, May 2019, 225-229

[4] D. Kimovski, R. Matha, G. Iuhasz, F. Marozzo, D. Petcu, R. Prodan, Autotuning of exascale applications with anomalies detection, Frontiers in Big Data, vol. 4, 101-115, 2021

[5] G. Iuhasz, D. Petcu, Monitoring of Exascale data processing, 2019 IEEE International Conference on Advanced Scientific Computing (ICASC), Romania, 2019, pp. 1-5.

[6] DIPET Consortium, Distributed Stream processing on Fog and Edge Systems via Transprecise Computing, CHIST-ERA programme, https://www.chistera.eu/projects/dipet

[7] COCO project team, Adaptivity in cloud-to-edge Continuum Computing , https://coco.hpc.uvt.ro/en/home/

[8] L. Cerd-Alabern, G. Iuhasz, G. Gemmi, Anomaly detection for fault detection in wireless community networks using machine learning, Computer Communications, 202, 191-203, 2023