h3-Open-BDEC: Innovative Software Platform for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning)

Kengo Nakajima
Information Technology Center
The University of Tokyo
Tokyo, Japan
nakajima@cc.u-tokyo.ac.jp

Takeshi Iwashita
Information Initiative Center
Hokkaido University
Sapporo, Japan
iwashita@iic.hokudai.ac.jp

Takashi Shimokawabe
Information Technology Center
The University of Tokyo
Tokyo, Japan
shimokawabe@cc.u-tokyo.ac.jp

Hiromichi Nagao
Earthquake Research Institute
The University of Tokyo
Tokyo, Japan
nagao@eri.u-tokyo.ac.jp

Takeshi Ogita
Division of Mathematical Sciences
Tokyo Woman's Christian University
Tokyo, Japan
ogita@lab.twcu.ac.jp

Takahiro Katagiri
Information Technology Center
Nagoya University
Nagoya, Japan
katagiri@cc.nagoya-u.ac.jp

Hisashi Yashiro
Ctr. for Global Environmental Research
Natl. Inst. for Environmental Studies
Tsukuba, Japan
yashiro.hisashi@nies.go.jp

Hiroya Matsuba
Center for Computational Science
RIKEN
Kobe, Japan
hiroya.matsuba@riken.jp

In this work, we propose an innovative method for computational science for sustainable promotion of scientific discovery by supercomputers in the Exascale Era by combining (Simulation + Data + Learning (S+D+L)), where ideas of data science and machine learning are introduced to computational science.

The BDEC system (Big Data & Extreme Computing), which is scheduled to be introduced to the Information Technology Center, the Tokyo University in 2021, is a Hierarchical, Hybrid, Heterogeneous (h3) system, which consists of computing nodes for computational science and those for data science/machine learning. In this study, we consider the BDEC as the platform for integration of (S+D+L), develop an innovative software platform "h3-Open-BDEC" for integration of (S+D+L), and evaluate the effects of integration of (S+D+L) on the BDEC. The h3-Open-BDEC (Fig.1) is designed for extracting the maximum performance of the supercomputers with minimum energy consumption focusing on (1) innovative method for numerical analysis with high-performance/high-reliability/power-saving based on the new principle of computing by adaptive precision, accuracy verification and automatic tuning, and (2) Hierarchical Data Driven Approach (hDDA) based on machine learning. In Data Driven Approach (DDA), technique of machine learning is introduced for predicting the results of simulations with different parameters. DDA generally requires a lot of simulations for generation of teaching data. We propose the hDDA, where simplified models for generating teaching data are constructed automatically by machine learning with Feature Detection, MOR, UQ, Sparse Modeling and AMR.

The h3-Open-BDEC is the first innovative software platform to realize integration of (S+D+L) on supercomputers in the Exascale Era, where computational scientists can achieve such integration without supports by other experts. Source codes and documents are open to public for various kinds of computational environments. This integration by h3-Open-BDEC enables significant reduction of computations and power consumptions, compared to those by conventional simulations.

This work is supported by Japanese Government from FY.2019 to FY.2023 (JSPS Grant-in-Aid for Scientific Research (S), P.I.: Kengo Nakajima) [1,2]

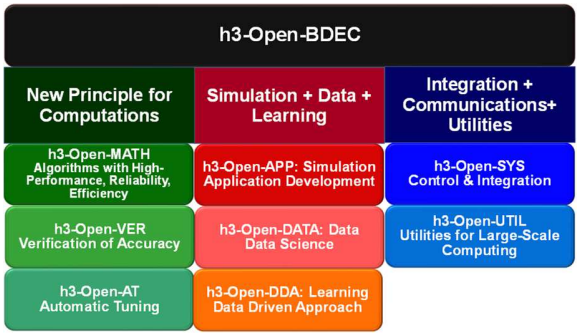

Fig.1 Overview of h3-Open-BDEC

REFERENCES

[1] https://kaken.nii.ac.jp/en/grant/KAKENHI-PROJECT-19H05662/
[2] http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC
[3] K. Nakajima, T. Furumura, T. Iwashita, T. Katagiri et al., ppOpen-HPC: Open Source Infrastructure for Development and Execution of Large-Scale Scientific Applications on Post-Peta-Scale Supercomputers with Automatic Tuning (AT), Mathematics for Industry 13, 15-35, Springer, 2015